THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Adaptive Variational Multiscale Methods

Axel Målqvist

CHALMERS | GÖTEBORG UNIVERSITY



Department of Mathematical Sciences
Chalmers University of Technology and Göteborg University
Göteborg, Sweden 2005

Adaptive Variational Multiscale Methods
Axel Målqvist
ISBN 91-7291-654-0

©Axel Målqvist, 2005.

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 2336
ISSN 0346-718X

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and Göteborg University
SE-412 96 Göteborg
Sweden
Telephone +46 (0)31 772 1000

Printed in Göteborg, Sweden 2005

Adaptive Variational Multiscale Methods

Axel Målqvist

Department of Mathematical Sciences
Chalmers University of Technology
Göteborg University

Abstract

In this thesis we present a new adaptive multiscale method for solving elliptic partial differential equations. The method is based on numerical solution of decoupled local fine scale problems on patches. Critical parameters such as fine and coarse scale mesh size and patch size are tuned automatically by an adaptive algorithm based on a posteriori error estimates. We extend the method to a mixed formulation of the Poisson equation and derive error estimates in this case as well.

We also present a framework for adaptivity based on a posteriori error estimates for multi-physics problems. We study a coupled flow and transport problem and derive an a posteriori error estimate for a linear functional by introducing two dual problems, one associated with the transport equation and one associated with the flow equation. We also apply this method to a model problem in oil reservoir simulation.

Keywords: finite element method, Galerkin, duality, a posteriori error estimation, adaptivity, variational multiscale method, mixed formulation, multi-physics, oil reservoir simulation

This thesis consists of the following papers:

Paper I: *Adaptive variational multiscale methods based on a posteriori error estimation: Energy norm estimates for elliptic problems,* (submitted) (with Mats G. Larson)

Paper II: *Adaptive variational multiscale methods based on a posteriori error estimation: Duality techniques for elliptic problems,* Lecture notes in: Computational Science and Engineering. B. Enquist, P. Löstedt, O. Runborg (eds.) Vol. 44, 181-193, 2005. (with Mats G. Larson)

Paper III: *A posteriori error estimates for mixed finite element approximations of elliptic problems,* (submitted) (with Mats G. Larson)

Paper IV: *A mixed adaptive variational multiscale method with applications in oil reservoir simulation,* (submitted)

Paper V: *Goal oriented adaptivity for coupled flow and transport problems with applications in oil reservoir simulation,* to appear in Math. Models Methods Appl. Sci. (with Mats G. Larson)

Related papers not included in the thesis:

Paper I: *A posteriori error analysis of the boundary penalty method,* (submitted) (with Mats G. Larson)

Paper II: *A posteriori error analysis of stabilized finite element approximations of the Helmholtz equation on unstructured grids,* (submitted) (with Mats G. Larson)

Paper III: *Adaptive variational multiscale methods based on a posteriori error estimation,* Proceedings of the 4th European Congress on Computational Methods in Applied Sciences and Engineering, P. Neittaanmäki, T. Rossi, S. Korotov, E. Onate, J. Periaux, and D. Knörzer (eds.), 2004. (with Mats G. Larson)

Paper IV: *Adaptive variational mutiscale methods for elliptic problems,* Proceedings of The Third M.I.T. Conference on Computational Fluid and Solid Mechanics, K. J. Bathe (ed.), 314-317, 2005. (with Mats G. Larson)

Contents

1 Introduction **1**
- 1.1 Thesis Objectives . . . 2
- 1.2 Main Results . . . 2
  - 1.2.1 A New Adaptive Variational Multiscale Method . . . 2
  - 1.2.2 A Framework for Adaptivity in Multi-Physics . . . 3
- 1.3 Future Work . . . 3

2 The Poisson Equation **3**
- 2.1 The Standard Formulation . . . 3
- 2.2 The Mixed Formulation . . . 4

3 The Finite Element Method **4**
- 3.1 The Weak Form . . . 4
- 3.2 Error Estimation . . . 5
- 3.3 Adaptive Algorithms . . . 5

4 Multiscale Methods **6**
- 4.1 The Variational Multiscale Method . . . 6
- 4.2 Adaptive Variational Multiscale Methods . . . 7
- 4.3 Adaptive Framework in Multi-Physics . . . 7

5 Oil Reservoir Simulation **8**

6 Summary of papers **8**
- 6.1 Paper I . . . 8
- 6.2 Paper II . . . 9
- 6.3 Paper III . . . 10
- 6.4 Paper IV . . . 10
- 6.5 Paper V . . . 10

Acknowledgements

First of all I would like to thank my supervisor Prof. Mats G. Larson. We started to work together on Helmholtz equation a winter break two and a half years ago and I think we have come a long way since then. Mats has a great view of our research field and his help and guidance has been invaluable.

Many thanks goes to Prof. Claes Johnson who gave me the opportunity to work as a PhD student in his group four years ago and who has supported my employment financially. Claes has through his brilliant mind opened doors not only for me but for many young computational mathematicians over the years.

I would like to thank the Department of Mathematical Sciences for financial support and for providing such a great working environment. In particular I would like to mention a few colleagues, my former roommate and good friend Georgios Foufas, my old class mates and companions on this journey Erik Broman and Elizabeth Wulcan, and from the Physics department the great Per Sundell. I also would like to thank my friends and former colleagues Fredrik Bengzon, Johan Hoffman, and Anders Logg for assisting with valuable comments on my work.

I would like thank my lovely parents Mabel and Bertil in Aneby and my dear brothers Gustav and Erik in Stockholm and Jönköping for giving me such a great support in life. Finally I would like to thank my dear Maria for her constant support, patience, and love.

Axel Målqvist
Göteborg, August 2005

1 Introduction

Multiscale and multi-physics problems are some of the greatest challenges in computational mathematics today. In all branches of the engineering sciences we encounter problems with several physical processes on several different scales. These problems often transforms into systems of partial differential equations with multiscale features in e.g. coefficients and these equations are very hard to solve within a reasonable tolerance on a single global mesh on a single processor.

Even if we for the moment neglect the multi-physical aspect and study one single equation with multiscale features we are still often forced to solve the problems by, in some way, splitting the equations into smaller subproblems that can be handled and then try to form a global solution out of these local approximations. Many methods using different versions of this idea has been developed over the last couple of decades and some of them have become important tools in engineering. Similarly, when solving multi-physics problems it is common to use optimized *black-box* solvers for each equation again breaking up the problem into simplified subproblems.

Both the idea of splitting up the multiscale problem into subproblems and using different solvers for different parts in a multi-physics computation introduces lots of new parameters to the equations. In an ordinary single mesh approach there are often just a few parameters such as the mesh size and size of the time steps. In multiscale and multi-physics solvers one often need to consider mesh sizes on all scales, size of subdomains for the local problems, boundary conditions for the local problems, different meshes for the different equations involved, different timesteps for different equation and so on. The more complicated the problems and methods get, the more parameters need to be set by the user.

This is not a trivial task since we often just have an idea of how these parameters relates in an asymptotic regime, and in many cases we are far from that in applications. It is obvious that there is a great need of guidance here. We need to make the solver understand when a certain mesh needs to be refined or when we need to solve some of our subproblems more accurately. We need clever solvers that can analyze already computed solutions and draw conclusions in order to improve the next version of the solution.

Such methods have been known for single mesh methods for a long time. In the finite element community they are referred to as adaptive algorithms. These algorithms are based on theoretical but computable bounds or estimates of the error committed by discretizing the problem. Adaptive algorithms work iteratively and uses earlier iterates of the solution to improve the next solution candidate.

We believe that such methods needs to be developed for multiscale and

multi-physics problems as well and we also believe that there is even more to gain from using adaptive methods in these fields.

1.1 Thesis Objectives

The main objectives of this thesis are to:

- Develop a new multiscale method for solving partial differential equations where error estimation and adaptivity is an integrated part of the method. We wish that the error estimate and adaptive algorithm of this method could serve as a framework for adaptivity in multiscale problems.

- Develop a framework for error estimation and adaptivity in multiphysics problems.

- Implement and test the algorithms on practically relevant test cases.

1.2 Main Results

1.2.1 A New Adaptive Variational Multiscale Method

- We develop a new adaptive multiscale method. We start from the variational multiscale formulation where the solution space is divided into a coarse and a fine part. We decouple the fine scale part into subproblems on patches and solve these local problems numerically using homogeneous Dirichlet boundary conditions. (Paper I)

- We derive an a posteriori error estimate for the energy norm and linear functionals of the error and present adaptive algorithms that automatically tunes the critical parameters of the method in order to obtain reliable and efficient approximations. We also implement the adaptive variational multiscale method and show that it works on basic test problems. (Paper I and Paper II)

- We derive an optimal order a posteriori error estimate in energy norm for mixed finite element methods using richer function spaces for the flux then the displacement. (Paper III)

- We extend the multiscale theory to a mixed formulation again with error estimates both for linear functionals and in energy norm, using the results from the three earlier papers and apply the mixed method to a model problem in oil reservoir simulation. (Paper IV)

1.2.2 A Framework for Adaptivity in Multi-Physics

- We present a framework for duality based a posteriori error estimation for multi-physical problems. (Paper V)
- We apply the theory to a coupled transport and flow problem that serves as a basic model problem in oil reservoir simulation.

1.3 Future Work

The development of adaptive algorithms based on a posteriori error estimates for multiscale and multiphysics problems has just started and we expect a rapid future development. There are thus many open problems and we formulate a few natural extensions that we have touched on during our work.

- Formulate the multiscale method for more then two scales and thereby be able to take on problems with more extreme scale separation.
- Make a comprehensive evaluation of how the method performs compared to other multiscale methods when solving for example the oil reservoir problem.
- Prove a priori error estimates for the multiscale method.
- Extend the multiscale method to time dependent problems.
- Extend the multiscale method to the transport equation and to even more challenging equations, for instance the Navier-Stokes equations.

2 The Poisson Equation

2.1 The Standard Formulation

Throughout this thesis we will mainly study the Poisson equation i.e.,

$$\begin{cases} -\triangle u = f & \text{in } \Omega, \\ u = 0 & \text{on } \Gamma, \end{cases} \tag{2.1}$$

where f are given data, Ω is a given domain with boundary Γ, and u is the unknown solution. The most characteristic feature of the equation is that it smoothes out rough parts of the load f which gives the solution higher regularity then the data. The equation models diffusion processes and appears in all branches of the engineering sciences such as heat transfer, structural mechanics, and electro magnetics.

2.2 The Mixed Formulation

By rewriting equation (2.1) we get the Poisson equation on mixed form

$$\begin{cases} \boldsymbol{\sigma} - \nabla u = 0 & \text{in } \Omega, \\ -\nabla \cdot \boldsymbol{\sigma} = f & \text{in } \Omega, \\ \boldsymbol{n} \cdot \boldsymbol{\sigma} = 0 & \text{on } \Gamma. \end{cases} \tag{2.2}$$

This is actually the same problem but this formulation is more common in applications such as flow in porous media where one is mainly interested in the flux $\boldsymbol{\sigma}$.

In this thesis we will work with both these versions of the Poisson equation but in the introduction we will focus on the standard formulation.

3 The Finite Element Method

The finite element method was developed by engineers in the 1950's and 1960's as an alternative to the finite difference methods for computing approximate solutions to differential equations. In the early days the method was mainly used in structural mechanics, see e.g. [25] for an overview. However, the finite element method also has a strong mathematical foundation in functional analysis, see [6]. The mathematical foundation provides the tools to derive analytical error estimates which can be used in a constructive way to improve the approximate solution. The method can also easily treat complex geometrical domains which makes it very useful in engineering applications.

3.1 The Weak Form

The first step in formulating a finite element method is to reformulate the problem on weak form. We introduce a set of test functions $v \in V = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma\}$, multiply equation (2.1) by v, integrate over the domain Ω and integrate by parts. The weak form reads: find $u \in V$ such that,

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx, \qquad \text{for all } v \in V. \tag{3.1}$$

The finite element method aims at approximating the solution to equation (3.1) rather then (2.1) which is different from the widely used finite difference method. Using the weak form gives access to lots of powerful tools from functional analysis which makes it possible to derive error estimates. We let V_h be a discrete approximation of V consisting of piecewise polynomials

defined on a mesh consisting of elements $\cup K = \mathcal{K}$ with diameter h_K, see Figure 1. The finite element methods reads: find $U \in V_h$ such that,

$$\int_\Omega \nabla U \cdot \nabla v \, dx = \int_\Omega f v \, dx, \qquad \text{for all } v \in V_h. \tag{3.2}$$

Since $V_h \subset V$, U will not be equal to u but hopefully a good approximation.

3.2 Error Estimation

There are two classes of finite element error estimates, *a priori* and *a posteriori*. The a priori estimate bounds the error $e = u - U$ in terms of data, u, and h while a posteriori estimates bounds e in terms of data, U, and h. In this thesis we will only consider a posteriori error bounds since they are computable once you have calculated U. A posteriori error estimates of the Poisson equation can be found in many books i.e. [11, 5, 17, 22]. We present such an estimate here in energy semi norm.

$$\|\nabla e\|^2 \leq C \sum_{K \in \mathcal{K}} \rho_K^2, \tag{3.3}$$

where $\rho_K^2 = h_K^2 \|f + \triangle U\|_K^2 + h_K \|[\boldsymbol{n} \cdot \nabla U]\|_{\partial K}^2$, ∂K is the boundary of K, $[\cdot]$ is the difference in function value over edges ∂K, and C is a constant independent of h.

3.3 Adaptive Algorithms

An adaptive algorithm is an iterative process that repeatedly solves the problem in richer and richer spaces V_h. We will now show how we can use equation (3.3) to formulate an adaptive algorithm:

1. Solve equation (3.2) on an initial mesh, see Figure 1 (left).
2. Calculate ρ_K from equation (3.3).
3. If $\sum_{K \in \mathcal{K}} \rho_K^2$ is sufficiently small stop, else refine the mesh according to ρ_K by splitting elements with big contributions to the error. Return to 1.

For a typical adaptive mesh see Figure 1 (right).

This sums up the section on adaptive finite element methods. There is obviously much more to say about this and we guide the reader to works as [17, 11] for an overview.

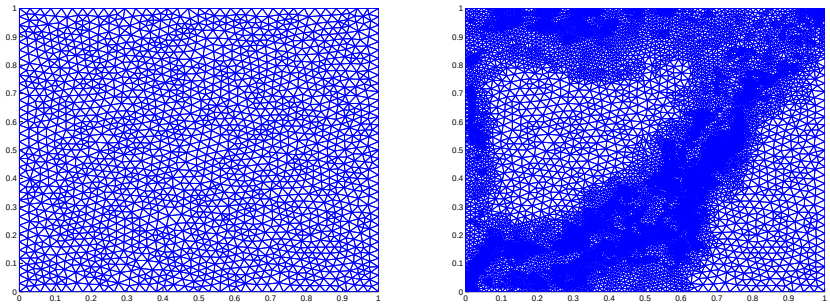


Figure 1: Quasi uniform mesh (left) and adaptively refined mesh (right).

4 Multiscale Methods

A theoretical foundation for multiscale methods can be found in Homogenization theory [8]. Homogenization deals mainly with problems where data has a periodic structure. For example we can add a space dependent coefficient to the Poisson equation (2.1),

$$\begin{cases} -\nabla \cdot a\nabla u = f & \text{in } \Omega, \\ u = 0 & \text{on } \Gamma, \end{cases} \tag{4.1}$$

where $a > 0$ and let $a = a(x/\epsilon)$ which means that a has a period of ϵ. In homogenization theory it is of interest to find an effective coefficient $\bar{a}$ with corresponding solution $\bar{U}$ that solves (4.1) in the limit when $\epsilon \to 0$. This means extreme scale separation that could for example be found in metals.

One multiscale method that has its roots in this theory and has got a lot of recognition is the multiscale finite element method presented in 1997 by Hou and Wu, [14]. In this method the standard finite element base functions are replaced by new base functions containing fine scale structure. These new base functions are calculated by solving local fine scale problems. Lately Aarnes and Lie have extended this method to the mixed setting (2.2) and applied it to oil reservoir problems, [2, 1].

4.1 The Variational Multiscale Method

Another important multiscale method is the variational multiscale method formulated by Hughes in 1995, see [15, 16]. This method is based on the idea of splitting the finite element spaces into a coarse and a fine part, $V = V_c \oplus V_f$. If we introduce this split in equation (3.1), rearrange terms and introduce the scalar product notation $(v, w) = \int_\Omega v\, w\, dx$ we get: find

$u_c \in V_c$ and $u_f \in V_f$ such that,

$$\begin{aligned} (\nabla u_c, \nabla v_c) + (\nabla u_f, \nabla v_c) &= (f, v_c) \quad \text{for all } v_c \in V_c, \\ (\nabla u_f, \nabla v_f) &= (f, v_f) - (\nabla u_c, \nabla v_f) \quad \text{for all } v_f \in V_f. \end{aligned} \tag{4.2}$$

The coarse space V_c is a standard finite element space but the fine space V_f is infinite dimensional. Hughes suggests that the fine scale solution u_f could be solved by analytical approximations on each coarse element in terms of u_c. After this is done u_c can be calculated directly in the coarse space V_c.

4.2 Adaptive Variational Multiscale Methods

The method we present in this thesis starts with the variational multiscale framework. We adopt the split between coarse and fine spaces but we solve the local problem numerically instead of analytically and we solve them on larger patches instead of on just one coarse element. This gives us two new parameters, namely the fine scale mesh size h and the patch size L, other then the coarse mesh size H.

Given this freedom it is important to use it wisely. We accomplish this by deriving an a posteriori error bound of the multiscale solution in terms of H, h, and L. We can now formulate an adaptive algorithm in the same spirit as in section 3.3, see [Paper I, Paper II] and [19, 24] for the Poisson equation on standard form and [Paper IV] for the Poisson on mixed form. These adaptive algorithms makes the method unique. Solving difficult problems with methods that lacks parameters that can be tuned or has parameters but no information on how they should be modified in order to improve the solution gives solutions that are difficult to evaluate and therefore rely on.

4.3 Adaptive Framework in Multi-Physics

In Paper V we study a different model problem,

$$\begin{cases} \dot{c} + \nabla \cdot (\boldsymbol{\sigma} c) - \epsilon \triangle c = g \quad \text{in } \Omega \times (0, T], \\ \qquad\qquad\qquad\qquad \boldsymbol{n} \cdot \nabla c = 0 \quad \text{on } \Gamma, \\ \qquad\qquad\qquad\qquad c = c_0 \quad \text{for } t = 0, \end{cases} \tag{4.3}$$

where g is given data, c_0 is initial data, and the flux $\boldsymbol{\sigma}$ is given as a solution to the following equation,

$$\begin{cases} \frac{1}{a}\boldsymbol{\sigma} - \nabla u = 0 \quad \text{in } \Omega, \\ -\nabla \cdot \boldsymbol{\sigma} = f \quad \text{in } \Omega, \\ \boldsymbol{n} \cdot \boldsymbol{\sigma} = 0 \quad \text{on } \Gamma. \end{cases} \tag{4.4}$$

This is a coupled transport and flow problem. The pressure equation (4.4) is however almost the same as the one we have studied earlier (2.2). We let

ϵ be very small since we want to minimize diffusion in our applications. To still get a stable solution we use streamline diffusion [13, 11, 17].

Again the a posteriori error estimate is the main theoretical result of the paper. Since we have a coupled set of PDE's we need to introduce a coupled set of dual problems to control the error in a given linear functional. It is often easy to decide in what sense the transport problem should be accurate. For example, a pointwise value could be of interest or an integrated quantity on an out flow boundary. It is not so easy to understand how this goal translates into the other equations that are present in a multi-physics problem. This is however solved by the error estimate we present in the last paper of this thesis. We also present an extensive numerical example on a simple petroleum reservoir model problem.

5 Oil Reservoir Simulation

Numerical simulation has become a more and more important complement to laboratory test and analytical models in oil reservoir simulation. To solve the full oil reservoir problem one would have to solve several coupled non linear partial differential equations, see for instance [2].

It is common to use various degrees of simplified versions of the oil reservoir problem, see for instance [9, 4, 1]. We have also chosen to study a simple model (4.3) just containing a flow problem and coupled with a transport problem. This model simulates the concentration of water in the reservoir after injecting water it a well in order to move the oil. Even though we do not consider the full complexity of the problem in this work we use proper data from the tenth SPE comparative project which gives us a similar fine scale structure as you would get from solving the full problem, see Figure 2. Since this problem both contains multiple scales and multi-physics it is very suitable to illustrate our ideas. We believe that there are many reasons for using adaptivity in oil reservoir simulation and we have taken a first step in that direction in this thesis.

6 Summary of papers

6.1 Paper I

Adaptive variational multiscale methods based on a posteriori error estimation: Energy norm estimates for elliptic problems, (submitted) (with Mats G. Larson)
The variational multiscale method (VMS) provides a general framework for construction of multiscale finite element methods. In this paper we propose

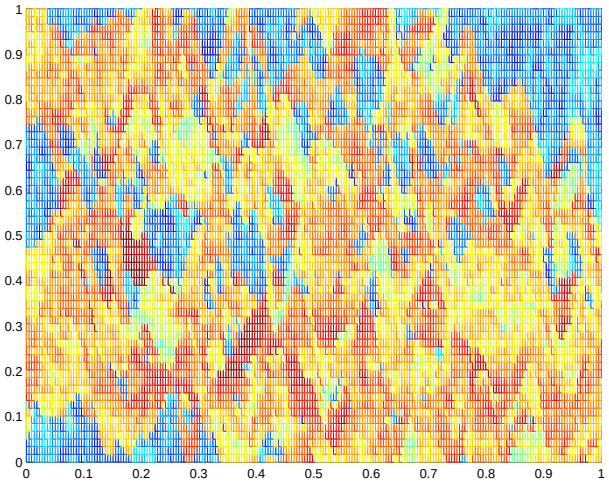


Figure 2: The permeability a used in the numerical example is taken from the Tenth SPE Comparative Solution Project. It is the bottom layer in the Upper Ness formation in log-scale

a method for parallel solution of the fine scale problem based on localized Dirichlet problems which are solved numerically. Next we present a posteriori error estimates for VMS which relates the error in energy norm to the discretization errors, resolution and size of patches in the localized problems, in the fine scale approximation. Based on the a posteriori error estimates we propose an adaptive VMS with automatic tuning of the critical parameters. We study elliptic second order partial differential equations with highly oscillating coefficients or localized singularities.

6.2 Paper II

Adaptive variational multiscale methods based on a posteriori error estimation: Duality techniques for elliptic problems Lecture notes in: Computational Science and Engineering. B. Enquist, P. Löstedt, O. Runborg (eds.) Vol. 44, 181-193, 2005. (with Mats G. Larson)
In applications it is very common to seek a specific functional of the solution rather than the solution itself. Such a functional may be a point value, derivative, or a mean value of the solution on a certain part of the domain. In this paper we extend the theory in Paper I to these kind of applications. We develop an a posteriori error estimate for a linear function of the error and present numerical tests where we for example decreases the error in a certain part of the domain by solving well refined local problems.

6.3 Paper III

A posteriori error estimates for mixed finite element approximations of elliptic problems, (submitted) (with Mats G. Larson)
When solving elliptic problem on mixed form using a richer space for the flux then the scalar variable it is known that a standard a posteriori error estimate will not give an optimal bound. One way to get around this problem is to replace the scalar variable in the method by a post processed improved version. In this paper we derive an a posteriori error estimate in energy norm that gives an optimal estimate for several commonly used elements.

6.4 Paper IV

A mixed adaptive variational multiscale method with applications in oil reservoir simulation, (submitted)
In this paper we extend the theory from Paper I and Paper II to the mixed formulation of the Poisson equation, equation (2.2). We solve local Neumann problems instead of local Dirichlet problems on the patches and use the results from Paper I, II, and III to derive an a posteriori error estimate in the energy norm and for a linear functional. This new adaptive multiscale method is applied to an elliptic problem from oil reservoir simulation. We show numerical examples where we illustrate how the adaptive algorithm automatically chooses where to put the most computational effort.

6.5 Paper V

Goal oriented adaptivity for coupled flow and transport problems with applications in oil reservoir simulation, to appear in Math. Models Methods Appl. Sci. (with Mats G. Larson)
We derive an error estimation framework for multi-physics problems. By associating a dual problem to each primal problem in a coupled set of equations we are able to minimize the error in a specific output quantity using different refinements for each individual equation. Standard a posteriori error estimators give information on which elements that needs to be refined in order to improve the solution. By using a set of dual problems we can also get error indicators that point out which equations that needs to be solved more efficiently in order to decrease the error. We apply the method to a coupled problem in oil reservoir simulation that contains a transport equation and an elliptic pressure equation. In the numerical experiments we use totally different elements and refinement strategies for the two equations.

References

[1] J. E. Aarnes, T. Gimse, and K.-A. Lie, *An introduction to the numerics of flow in porous media using Matlab,* Geometrical Modelling Numerical Simulation and Optimization, Industrial Mathematics at SINTEF. Springer Verlag, 2005.

[2] J. E. Aarnes, K.-A. Lie, and A. B. Rustad, *Modelling of multiscale structures in flow simulations for petroleum reservoirs,* submitted.

[3] R. A. Adams, *Sobolev spaces,* volume 65 of Pure and Applied Mathematics, Academic Press, New York, 1975.

[4] T. Arbogast and S. L. Bryant, *A two-scale numerical subgrid technique for waterflood simulations,* SPE J., 2002.

[5] W. Bangerth and R. Rannacher *Adaptive finite element methods for differential equations,* Birkhäuser, 2003.

[6] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods,* Springer Verlag, 1994.

[7] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods,* Springer Verlag, 1991.

[8] D. Cioranescu and P. Donato, *An introduction to homogenization,* Oxford University Press, 2000.

[9] C. Dawson, S. Sun, and M. F. Wheeler, *Compatible algorithms for coupled flow and transport,* Comput. Methods Appl. Mech. Engrg., 193, 2565-2580, 2004.

[10] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Introduction to adaptive methods for differential equations,* Acta Numerica, 105-158, 1995.

[11] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, *Computational differential equations,* Studentlitteratur, 1996.

[12] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems I: A linear model problem,* SIAM J. Numer. Anal., 28, 43-77, 1991.

[13] K. Eriksson and C. Johnson, *Adaptive streamline diffusion finite element methods for stationary convection-diffusion problems,* Math. Comp., 60, 167-188, 1993.

[14] T. Y. Hou and X.-H. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media,* J. Comput. Phys., 134, 169-189, 1997.

[15] T. J.R. Hughes, *Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods,* Comput. Methods Appl. Mech. Engrg. 127, 387-401, 1995.

[16] T. J. R. Hughes, G. R. Feijóo, L. Mazzei and Jean-Baptiste Quincy, *The variational multiscale method - a paradigm for computational mechanics,* Comput. Methods Appl. Mech. Engrg., 166, 3-24, 1998.

[17] C. Johnson, *Numerical solution of partial differential equations by the finite element method,* Studentlitteratur, 1987.

[18] M. G. Larson and A. Målqvist, *Adaptive variational multiscale methods based on a posteriori error estimation: Duality techniques for elliptic problems,* Lecture notes in: Computational Science and Engineering. B. Enquist, P. Löstedt, O. Runborg (eds.) Vol. 44, 181-193, 2005.

[19] M. G. Larson and A. Målqvist, *Adaptive variational multiscale methods based on a posteriori error estimation,* Proceedings of the 4th European Congress on Computational Methods in Applied Sciences and Engineering, P. Neittaanmäki, T. Rossi, S. Korotov, E. Onate, J. Periaux, and D. Knörzer (eds.), 2004.

[20] M. G. Larson and A. Målqvist, *Adaptive variational multiscale methods based on a posteriori error estimation: Energy norm estimates for elliptic problems,* submitted.

[21] M. G. Larson and A. Målqvist, *Adaptive variational mutiscale methods for elliptic problems,* Proceedings of The Third M.I.T. Conference on Computational Fluid and Solid Mechanics, K. J. Bathe (ed.), 314-317, 2005.

[22] S. Larsson and V. Thomée, *Partial differential equations with numerical methods,* Springer Verlag, 2005.

[23] A. Målqvist, *A mixed adaptive variational multiscale method with applications in oil reservoir simulation,* submitted.

[24] A. Målqvist and M. G. Larson, *Goal oriented adaptivity for coupled flow and transport problems with applications in oil reservoir simulation,* to appear in Math. Models Methods Appl. Sci.

[25] C. O. Zienkiewicz, *The birth of the finite element method and of computational mechanics,* Internat. J. Numer. Methods Engrg., 60, 3-10, 2004.

Adaptive Variational Multiscale Methods Based on A Posteriori Error Estimation: Energy Norm Estimates for Elliptic Problems

Mats G. Larson* and Axel Målqvist†

August 19, 2005

Abstract

We develop a new adaptive multiscale finite element method using the variational multiscale framework together with a systematic technique for approximation of the fine scale part of the solution. The fine scale is approximated by a sum of solutions to decoupled localized problems, which are solved numerically on a fine grid partition of a patch of coarse grid elements. The sizes of the patches of elements may be increased to control the error caused by localization. We derive an a posteriori error estimate in the energy norm which captures the dependency of the crucial discretization parameters: the coarse grid mesh size, the fine grid mesh size, and the sizes of the patches. Based on the a posteriori error estimate we present an adaptive algorithm that automatically tunes these parameters. Finally, we show how the method works in practice by presenting various numerical examples.

1 Introduction

The application of multiscale problems are numerous. They appear in all branches of the engineering sciences, for instance, composite materials, flow in porous media, fluid mechanics, and quantum physics. A common feature for all these applications is that they are very computationally challenging and often impossible to solve to an acceptable tolerance with standard one mesh methods. We thus need to develop new methods which are based

*Corresponding author, Professor, Department of Mathematics, Umeå University, S-90187 Umeå, mats.larson@math.umu.se.

†Graduate Research Assistant, Department of Mathematics, Chalmers University of Technology, S-41296 Göteborg, axel@math.chalmers.se.

on a combination of global and localized computations, so called multiscale methods. Further to guarantee accuracy we need to develop error estimates and adaptive algorithms. In this paper we take a first step in this direction.

Previous work. The Variational Multiscale (VMS) method serves as a general framework for the solution of multiscale problems, see [9, 10]. The idea is to decompose the solution into fine and coarse scale contributions, solve the fine scale equations in terms of the coarse scale residual, and finally eliminate the fine scale solution from the coarse scale equation. This procedure leads to a modified coarse scale equation where the modification accounts for the effect of fine scale behavior on the coarse scales. In several works various ways of analytical modeling are investigated often based on bubbles or element Green's functions, see Hughes [9], Oberai and Pinsky, [15], and Arbogast [1]. In [7, 8] Hou et.al. present a different approach. Here the fine scale equations are solved numerically on a finer mesh. The fine scale equations are solved inside coarse elements and are thus totally decoupled. To reduce the effect from the boundary conditions forced on the patches Hou and Wu also present a method using slightly larger patches called the over-sampling method, see [8].

In the adaptive variational multiscale method (AVMS), first introduced by Larson and Målqvist [13, 12], the fine scale equations are decoupled and solved approximately on patches. In [12] an a posteriori error estimate is presented for control of a linear functional of the error. The method is adaptive in the sense that both the subdomains where local problems are solved and the resolution in these local problems are chosen automatically.

New contributions. The focus of this paper is to present the AVMS method and to derive an a posteriori error estimate in the energy norm. The basic idea of AVMS is to split the fine scale residual into localized contributions using a partition of unity and solving corresponding decoupled localized problems on patches with homogeneous Dirichlet boundary conditions. The fine scale solution is approximated by the sum $U_f = \sum_i U_{f,i}$ of the solutions $U_{f,i}$ to the localized problems associated with coarse node i. The accuracy of U_f depends on the fine scale mesh size h and the size of the patches. We note that the fine scale computation is naturally parallel.

To optimize performance we want to construct an adaptive algorithm for automatic control of the coarse mesh size H, the fine mesh size h, and the size of patches. The algorithm is based on the following a posteriori estimate of the error $e = u - U_c - U_f$ in the energy norm for the Poisson

equation with variable coefficient a:

$$\|e\|_a^2 \leq C \sum_{i \in \mathcal{C}} \|H\mathcal{R}(U_c)\|_{\omega_i}^2 \|\frac{1}{\sqrt{a}}\|_{L^\infty(\omega_i)}^2 \tag{1.1}$$
$$+ C \sum_{i \in \mathcal{F}} \left(\|\sqrt{H}\Sigma(U_{f,i})\|_{\partial\omega_i}^2 + \|h\mathcal{R}_i(U_{f,i})\|_{\omega_i}^2 \right) \|\frac{1}{\sqrt{a}}\|_{L^\infty(\omega_i)}^2,$$

where

$$(-\Sigma(U_{f,i}), v_f)_{\partial\omega_i} = (f + \nabla \cdot a \nabla U_c, \varphi_i v_f)_{\omega_i} - a(U_{f,i}, v_f)_{\omega_i}, \text{ for all } v_f \in V_f^h(\bar{\omega_i}), \tag{1.2}$$

$\mathcal{C}$ refers to nodes where no local problems have been solved, $\mathcal{F}$ to nodes where local problems are solved, U_c is the coarse scale solution, $U = U_c + U_f$, $\{\varphi_i\}_{i \in \mathcal{C} \cup \mathcal{F}}$ is a partition of unity, $\mathcal{R}(U)$ is a computable bound of the residual $f + \nabla \cdot a \nabla U$, $\mathcal{R}_i(U_{f,i})$ is a bound of the fine scale residual $\varphi_i(f + \nabla \cdot a \nabla U_c) + \nabla \cdot a \nabla U_{f,i}$, $\Sigma(U_{f,i})$ is related to the normal derivative of the fine scale solution $U_{f,i}$ and measures the effect of restriction to patches, and $V_f^h(\bar{\omega_i})$ is a finite element space on the fine scale local problem i. If no fine scale equations are solved we obtain the first term in the estimate; the first part of the second sum measures the effect of restriction to patches; and finally the second part measures the influence of the fine scale mesh parameter h.

The framework of AVMS is fairly general and may be extended to other types of multiscale methods, for instance, based on localized Neumann problems.

Outline. The remainder of the paper is organized as follows: in Section 2 we introduce the model problem and the adaptive variational multiscale formulation; in Section 3 we present and prove a posteriori error estimates; in Section 4 we study the special case of periodic coefficients; in Section 5 we present an adaptive algorithm based on the a posteriori error estimate; and, finally, in Section 6 we present illustrating numerical results.

2 The Variational Multiscale Method

2.1 Model Problem

We study the Poisson equation with variable coefficients and homogeneous Dirichlet boundary conditions: find $u \in H_0^1(\Omega)$ such that

$$-\nabla \cdot a \nabla u = f \quad \text{in } \Omega, \tag{2.1}$$
$$u = 0 \quad \text{on } \Gamma, \tag{2.2}$$

where Ω is a polygonal domain in $\mathbf{R}^d$, $d = 1, 2$, or 3 with boundary Γ, $f \in L^2(\Omega)$, and $a \in L^\infty(\Omega)$ such that $a(x) > 0$ for all $x \in \Omega$. The variational form of (2.1) reads: find $u \in \mathcal{V} = H_0^1(\Omega)$ such that

$$a(u, v) = (f, v) \quad \text{for all } v \in \mathcal{V}, \tag{2.3}$$

with the bilinear form

$$a(u, v) = (a\nabla u, \nabla v), \tag{2.4}$$

for all $u, v \in \mathcal{V}$.

2.2 The Variational Multiscale Method

We focus on two scales and employ the variational multiscale scale formulation, see Hughes [9, 10] for an overview. We choose two spaces $\mathcal{V}_c \subset \mathcal{V}$ and $\mathcal{V}_f \subset \mathcal{V}$ such that

$$\mathcal{V} = \mathcal{V}_c \oplus \mathcal{V}_f, \tag{2.5}$$

where $\mathcal{V}_c$ is associated with the coarse scale and $\mathcal{V}_f$ is associated with the fine scale. Introducing these spaces in (2.3) gives us the following weak formulation: find $u_c \in \mathcal{V}_c$ and $u_f \in \mathcal{V}_f$ such that

$$\begin{aligned} a(u_c, v_c) + a(u_f, v_c) &= (f, v_c) \quad \text{for all } v_c \in \mathcal{V}_c, \\ a(u_c, v_f) + a(u_f, v_f) &= (f, v_f) \quad \text{for all } v_f \in \mathcal{V}_f. \end{aligned} \tag{2.6}$$

We let $R : \mathcal{V} \to \mathcal{V}'$ denote the residual defined by

$$(R(v), w) = (f, w) - a(v, w) \quad \text{for all } w \in \mathcal{V}. \tag{2.7}$$

The fine scale equation now takes the form: find $u_f \in \mathcal{V}_f$ such that

$$a(u_f, v_f) = (R(u_c), v_f) \quad \text{for all } v_f \in \mathcal{V}_f. \tag{2.8}$$

Thus the fine scale solution is driven by the residual of the coarse scale solution. Denoting the solution u_f to (2.8) by $u_f = \mathcal{T}R(u_c)$ we get the modified coarse scale problem

$$a(u_c, v_c) + a(\mathcal{T}R(u_c), v_c) = (f, v_c) \quad \text{for all } v_c \in \mathcal{V}_c. \tag{2.9}$$

Here the second term on the left hand side accounts for the effects of fine scales on the coarse scales.

In terms of matrices this gives us a modified stiffness matrix and a modified right hand side since $a(\mathcal{T}R(\phi_i), \phi_j) = T_{ij} + d_j$ for some matrix T and vector d. Note that $(R(v), w)$ defined in equation (2.7) is affine in v. If we

denote the standard Galerkin stiffness matrix by A and the right hand side by b we would get $AU_G = b$ for the standard Galerkin but

$$(A + T)U_c = b - d, \tag{2.10}$$

for the modified version.

Another approach is to solve equation (2.6) iteratively with the Galerkin solution as an initial guess.

2.3 Approximation of Fine Scale Equations Based on Localized Dirichlet Problems

We use the method described in our earlier work [12, 13] for the approximate solution of the fine scale equations. The idea is to decouple the fine scale equations by including a partition of unity in the right hand side and then solve the resulting problems on patches. We start with some preliminary notations.

We introduce a partition $\mathcal{K} = \{K\}$ of the domain Ω into coarse shape regular elements K of diameter H_K and we let $\mathcal{N}$ be the set of coarse nodes. Further we let $\mathcal{V}_c$ be the space of continuous piecewise polynomials of degree p defined on $\mathcal{K}$.

We let $u_f = \sum_{i\in\mathcal{N}} u_{f,i}$ where

$$a(u_{f,i}, v_f) = (\varphi_i R(u_c), v_f) \quad \text{for all } v_f \in \mathcal{V}_f, \tag{2.11}$$

and $\{\varphi_i\}_{i\in\mathcal{N}}$ is a partition of unity e. g. the set of Lagrange basis functions in $\mathcal{V}_c$, be the solution to the decoupled fine scale equations. We note that the right hand side has the same support as φ_i and a small support compared to Ω.

We introduce this expansion of u_f in the right hand side of the fine scale equation (2.6) and get: find $u_c \in \mathcal{V}_c$ and $u_f = \sum_{i\in\mathcal{N}} u_{f,i} \in \mathcal{V}_f$ such that

$$\begin{aligned} a(u_c, v_c) + a(u_f, v_c) &= (f, v_c) \quad \text{for all } v_c \in \mathcal{V}_c, \\ a(u_{f,i}, v_f) &= (\varphi_i R(u_c), v_f) \quad \text{for all } v_f \in \mathcal{V}_f \text{ and } i \in \mathcal{N}. \end{aligned} \tag{2.12}$$

The next step is to solve the fine scale equations approximately. For each element φ_i in the partition of unity we associate a domain ω_i on which we solve Dirichlet problems. The local domain ω_i contains the support of the element in the partition of unity φ_i and will be chosen large enough to give a good approximate solution. The quality of the solution is controlled by error estimates. We now define the local finite element space $\mathcal{V}_f^h(\omega_i)$ associated with node i. We refine the coarse mesh on the patch ω_i and let $\mathcal{V}_f^h(\omega_i)$ be the fine part of the nodal hierarchical basis on this mesh that fulfills homogeneous Dirichlet boundary conditions on the boundary of the

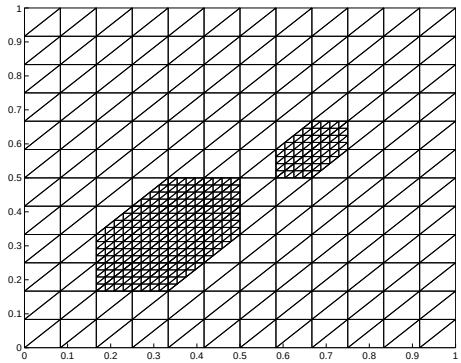


Figure 1: Two (left) and one (right) layer stars.

patch ω_i. This means that functions in $\mathcal{V}_f^h(\omega_i)$ are continuous piecewise polynomials of degree p on the fine scale equal to zero in the coarse nodes. We let h denote the fine scale mesh parameter associated with these spaces. In Figure 1 we show an examples of patches with fine meshes and the coarse mesh.

The resulting method reads: find $U_c \in \mathcal{V}_c$ and $U_f = \sum_{i\in\mathcal{N}} U_{f,i}$ where $U_{f,i} \in \mathcal{V}_f^h(\omega_i)$ such that

$$\begin{aligned} a(U_c, v_c) + a(U_f, v_c) &= (f, v_c) \quad \text{for all } v_c \in \mathcal{V}_c, \\ a(U_{f,i}, v_f) &= (\varphi_i R(U_c), v_f) \quad \text{for all } v_f \in \mathcal{V}_f^h(\omega_i) \text{ and } i \in \mathcal{N}. \end{aligned} \tag{2.13}$$

Since the functions in the local finite element spaces $\mathcal{V}_f^h(\omega_i)$ are equal to zero on $\partial\omega_i$, U_f and therefore U will be continuous.

Remark 2.1 For problems with multiscale phenomena on a part of the domain it is not necessary to solve local problems for all coarse nodes. We let $\mathcal{C} \subset \mathcal{N}$ refer to nodes where no local problems are solved and $\mathcal{F} \subset \mathcal{N}$ refer to nodes where local problems are solved. Obviously $\mathcal{C} \cup \mathcal{F} = \mathcal{N}$. We let $U_{f,i} = 0$ for $i \in \mathcal{C}$.

Remark 2.2 The choice of the subdomains ω_i is crucial for the method. We introduce a notation for mesh stars of many layers of coarse elements recursively in the following way. Let S_1^i be the support of the coarse scale Lagrangian base function φ_i in node i. The extended mesh star $S_L^i = \cup_{x_j \in S_{L-1}^i} S_1^j$ for $L > 1$, where $\varphi_j(x_j) = 1$, nodes on the boundary are included. We refer to L as the numbers of layers, see Figure 1.

Remark 2.3 Note that the construction of the fine scale local problems are

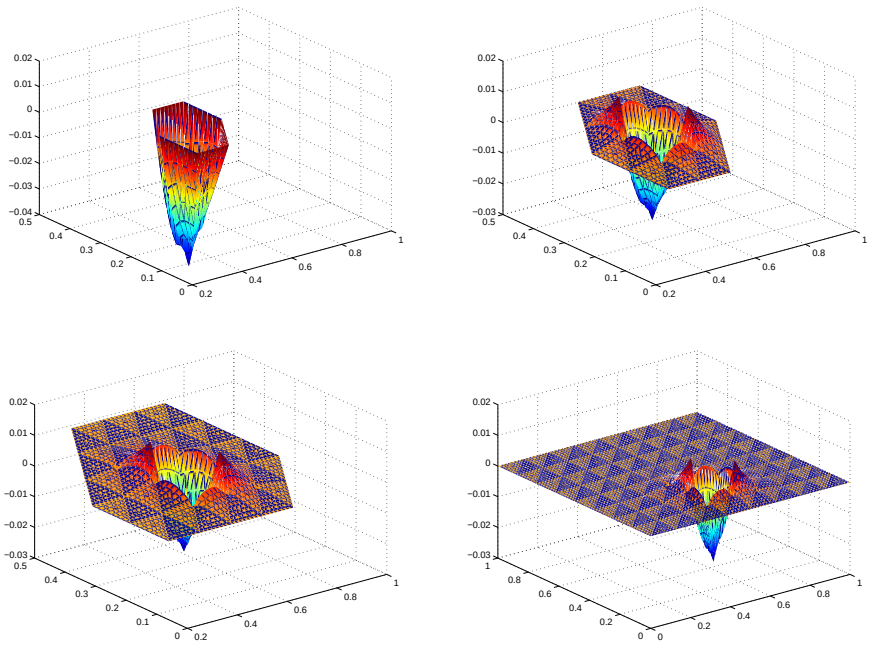


Figure 2: A typical localized solution $U_{f,i}$ of the fine scale equations in a smooth region using one, two, three layer stars, and the entire domain.

inspired by energy norm a posteriori error estimates based on local Dirichlet problems on stars. See [2] and [14] for further details on such estimates.

To understand how the localized solution $U_{f,i}$ behaves when the domain ω_i is increased we plot different solutions $U_{f,i}$ in a smooth region of the solution u in Figure 2. Since $U_{f,i}$ is solved in the slice space $\mathcal{V}_f$ and since the right hand side of the fine scale equations of (2.13) has the same support as φ_i, $U_{f,i}$ will decay rapidly towards the boundary of ω_i, this can also be seen in Figure 2. We note that one layer stars appears to give bad accuracy while two and more layer stars captures the main features of the exact fine scale solution $u_{f,i}$.

If we let $\omega_i = \Omega$ we will get the same solution as if we solve the global problem on the fine mesh, given that h in constant between the patches. However, to determine a suitable number of layers we develop error estimates.

3 A Posteriori Error Estimate in the Energy Norm

We start by introducing notations for bounds of the residual. Let $\mathcal{R}(U)$ be a bound of the residual defined in the following way, see [6]:

$$\mathcal{R}(U) = |f + \nabla \cdot a\nabla U| + \frac{1}{2}\max_{\partial K\setminus\Gamma} H_K^{-1}|[a\partial_n U]| \quad \text{on } K \in \mathcal{K}, \tag{3.1}$$

where $\mathcal{K}$ is the set of elements in the mesh and $[\cdot]$ is the difference in function value over the current interior edge. We note that $|(R(U), v)| \leq \|H^s\mathcal{R}(U)\|\|H^{-s}v\|$ for $s \in \mathbf{R}$. We define $\mathcal{R}_i(U_{f,i})$ in the same way as $\mathcal{R}(U)$ on the local mesh but with U replaced by $U_{f,i}$, H by h, and f by $\varphi_i\mathcal{R}(U_c)$.

We also define a new space on the patches ω_i. Let $\mathcal{V}_f^h(\bar{\omega}_i)$ be the fine scale part of the hierarchical space of piecewise polynomials of degree p defined on the fine mesh on patch ω_i. This space is identical to $\mathcal{V}_f^h(\omega_i)$ with the difference that $\mathcal{V}_f^h(\bar{\omega}_i)$ is not necessarily zero on the boundary $\partial\omega_i$. This means that $\mathcal{V}_f^h(\omega_i) \subset \mathcal{V}_f^h(\bar{\omega}_i)$. We also introduce the space $\mathcal{V}_f^h(\partial\omega_i)$ which is the restriction of $\mathcal{V}_f^h(\bar{\omega}_i)$ on the boundary $\partial\omega_i$. For a given function $v \in \mathcal{V}_f^h(\bar{\omega}_i)$ we have that v restricted to the boundary is in $\mathcal{V}_f^h(\partial\omega_i)$.

We derive an error estimate involving both the coarse scale error $e_c = u_c - U_c$ and the fine scale error $e_f = \sum_{i\in\mathcal{N}} e_{f,i} := \sum_{i\in\mathcal{N}}(u_{f,i} - U_{f,i})$ that arises from using our finite element method (2.13).

If we subtract the coarse part of equation (2.13) from the coarse part of equation (2.12) we get the Galerkin orthogonality,

$$a(e_c, v_c) + a(e_f, v_c) = 0 \quad \text{for all } v_c \in \mathcal{V}_c. \tag{3.2}$$

The same argument on the fine scale equation gives for $i \in \mathcal{F}$,

$$a(e_{f,i}, v_f) = -a(e_c, \varphi_i v_f), \quad \text{for all } v_f \in \mathcal{V}_f^h(\omega_i). \tag{3.3}$$

We state the following estimate for the error in the energy norm, $\|e\|_a = a(e,e)^{1/2}$.

Theorem 3.1 *Let $a \in L^\infty(\Omega)$ and $f \in L^2(\Omega)$. Then the error $e = u - U$ where u is the exact solution to (2.1) and $U = U_c + \sum_{i\in\mathcal{N}} U_{f,i}$ is the approximate solution defined by (2.13) satisfies the estimate*

$$\begin{aligned}\|e\|_a^2 \leq C\sum_{i\in\mathcal{C}} \|H\mathcal{R}(U_c)\|_{\omega_i}^2 \|\frac{1}{\sqrt{a}}\|_{L^\infty(\omega_i)}^2 \\ + C\sum_{i\in\mathcal{F}} \left(\|\sqrt{H}\Sigma(U_{f,i})\|_{\partial\omega_i}^2 + \|h\mathcal{R}_i(U_{f,i})\|_{\omega_i}^2\right)\|\frac{1}{\sqrt{a}}\|_{L^\infty(\omega_i)}^2,\end{aligned} \tag{3.4}$$

where $\Sigma(U_{f,i}) \in \mathcal{V}_f^h(\partial\omega_i)$ *is defined by*

$$(-\Sigma(U_{f,i}), v_f)_{\partial\omega_i} = (\varphi_i R(U_c), v_f)_{\omega_i} - a(U_{f,i}, v_f)_{\omega_i}, \quad \textit{for all } v_f \in V_f^h(\bar{\omega_i}). \tag{3.5}$$

Proof. We use the error equation (3.2) with v_c as the Scott-Zhang interpolant $\pi_c e$ onto the coarse space $\mathcal{V}_c$, see [3], to get,

$$\begin{aligned}
\|e\|_a^2 &= a(e,e) && (3.6)\\
&= a(e, e-\pi_c e) && (3.7)\\
&= a(u-U_c, e-\pi_c e) - a(U_f, e-\pi_c e) && (3.8)\\
&= (R(U_c), e-\pi_c e) - a(U_f, e-\pi_c e) && (3.9)\\
&= \sum_{i\in\mathcal{C}} (\varphi_i R(U_c), e-\pi_c e) && (3.10)\\
&\quad + \sum_{i\in\mathcal{F}} (\varphi_i R(U_c), e-\pi_c e) - a(U_{f,i}, e-\pi_c e)\\
&= \sum_{i\in\mathcal{C}} (\varphi_i R(U_c), e-\pi_c e) && (3.11)\\
&\quad + \sum_{i\in\mathcal{F}} (\varphi_i R(U_c), \pi_{f,i}(e-\pi_c e)) - a(U_{f,i}, \pi_{f,i}(e-\pi_c e))\\
&\quad + \sum_{i\in\mathcal{F}} (\varphi_i R(U_c), e-\pi_c e - \pi_{f,i}(e-\pi_c e))\\
&\quad\quad - \sum_{i\in\mathcal{F}} a(U_{f,i}, e-\pi_c e - \pi_{f,i}(e-\pi_c e))\\
&= \mathrm{I} + \mathrm{II} + \mathrm{III}, && (3.12)
\end{aligned}$$

where $\pi_{f,i}$ is the Scott-Zhang interpolant onto $\mathcal{V}_f^h(\bar{\omega_i})$. We start by estimating the first term of equation (3.12), I. From interpolation theory [3] we have,

$$\begin{aligned}
\sum_{i\in\mathcal{C}} (\varphi_i R(U_c), e-\pi_c e) &\le \sum_{i\in\mathcal{C}} \|\varphi_i \mathcal{R}(U_c)\|_{\omega_i} \|e-\pi_c e\|_{\omega_i} && (3.13)\\
&\le C \sum_{i\in\mathcal{C}} \|H\mathcal{R}(U_c)\|_{\omega_i} \|\nabla e\|_{\omega_i}. && (3.14)
\end{aligned}$$

Next we turn our attention to the second term of equation (3.12), II.

We introduce $\Sigma(U_{f,i}) \in \mathcal{V}_f^h(\partial\omega_i)$ as the piecewise polynomial defined on $\partial\omega_i$ that uniquely solves,

$$(-\Sigma(U_{f,i}), v_f)_{\partial\omega_i} = (R(U_c), \varphi_i v_f)_{\omega_i} - a(U_{f,i}, v_f)_{\omega_i}, \quad \text{for all } v_f \in V_f^h(\bar{\omega_i}). \tag{3.15}$$

With this definition we get the following estimate for the second term,

$$\mathrm{II} = \sum_{i\in\mathcal{F}} (-\Sigma(U_{f,i}), \pi_{f,i}(e-\pi_c e))_{\partial\omega_i} \tag{3.16}$$

$$\leq \sum_{i\in\mathcal{F}} \|\sqrt{H}\Sigma(U_{f,i})\|_{\partial\omega_i} \|\frac{1}{\sqrt{H}}\pi_{f,i}(e-\pi_c e)\|_{\partial\omega_i}. \tag{3.17}$$

We use the the following trace inequality from [3],

$$\|\pi_{f,i}(e-\pi_c e)\|^2_{\partial\omega_i} \leq C\left(\frac{1}{H}\|\pi_{f,i}(e-\pi_c e)\|^2_{\omega_i} + H\|\nabla\pi_{f,i}(e-\pi_c e)\|^2_{\omega_i}\right). \tag{3.18}$$

Next we use that the Scott-Zhang interpolant is both L^2 and H^1 stable on shape-regular meshes from [5, 4] to get,

$$\|\pi_{f,i}(e-\pi_c e)\|^2_{\partial\omega_i} \leq C\left(\frac{1}{H}\|e-\pi_c e\|^2_{\omega_i} + H\|\nabla(e-\pi_c e)\|^2_{\omega_i}\right) \tag{3.19}$$

$$\leq CH\|\nabla e\|^2_{\omega_i}. \tag{3.20}$$

We conclude

$$\mathrm{II} \leq C\sum_{i\in\mathcal{F}} \|\sqrt{H}\Sigma(U_{f,i})\|_{\partial\omega_i}\|\nabla e\|_{\omega_i}. \tag{3.21}$$

We now take on the third term in equation (3.12), $\sum_{i\in\mathcal{F}}(\varphi_i R(U_c), e-\pi_c e - \pi_{f,i}(e-\pi_c e)) - a(U_{f,i}, e-\pi_c e-\pi_{f,i}(e-\pi_c e))$,

$$\mathrm{III} \leq C\sum_{i\in\mathcal{F}} \|h\mathcal{R}_i(U_{f,i})\|_{\omega_i}\|\nabla(e-\pi_c e)\|_{\omega_i} \tag{3.22}$$

$$\leq C\sum_{i\in\mathcal{F}} \|h\mathcal{R}_i(U_{f,i})\|_{\omega_i}\|\nabla e\|_{\omega_i}. \tag{3.23}$$

We need to make the following simple observation,

$$\|\nabla e\|_{\omega_i} \leq \|\frac{1}{\sqrt{a}}\|_{L^\infty(\omega_i)}\|\sqrt{a}\nabla e\|_{\omega_i}, \tag{3.24}$$

by Hölder's inequality. We go back to equation (3.6) and use the estimates

of the three terms together with equation (3.24)

$$\begin{aligned}
\|e\|_a^2 &\leq \sum_{i\in\mathcal{C}}(\varphi_i R(U_c), e-\pi_c e) \qquad (3.25)\\
&\quad + \sum_{i\in\mathcal{F}}(\varphi_i R(U_c), \pi_{f,i}(e-\pi_c e)) - a(U_{f,i}, \pi_{f,i}(e-\pi_c e))\\
&\quad + \sum_{i\in\mathcal{F}}(\varphi_i R(U_c), e-\pi_c e-\pi_{f,i}(e-\pi_c e))\\
&\quad - \sum_{i\in\mathcal{F}} a(U_{f,i}, e-\pi_c e-\pi_{f,i}(e-\pi_c e))\\
&\leq C\sum_{i\in\mathcal{C}} \|H\mathcal{R}(U_c)\|_{\omega_i}\|\nabla e\|_{\omega_i} \qquad (3.26)\\
&\quad + C\sum_{i\in\mathcal{F}} \|\sqrt{H}\Sigma(U_{f,i})\|_{\partial\omega_i}\|\nabla e\|_{\omega_i}\\
&\quad + C\sum_{i\in\mathcal{F}} \|h\mathcal{R}_i(U_{f,i})\|_{\omega_i}\|\nabla e\|_{\omega_i}\\
&\leq C\left(\sum_{i\in\mathcal{C}} \|H\mathcal{R}(U_c)\|_{\omega_i}^2 \|\frac{1}{\sqrt{a}}\|_{L^\infty(\omega_i)}^2\right)^{1/2}\|e\|_a \qquad (3.27)\\
&\quad + C\left(\sum_{i\in\mathcal{F}} \|\sqrt{H}\Sigma(U_{f,i})\|_{\partial\omega_i}^2 \|\frac{1}{\sqrt{a}}\|_{L^\infty(\omega_i)}^2\right)^{1/2}\|e\|_a\\
&\quad + C\left(\sum_{i\in\mathcal{F}} \|h\mathcal{R}_i(U_{f,i})\|_{\omega_i}^2 \|\frac{1}{\sqrt{a}}\|_{L^\infty(\omega_i)}^2\right)^{1/2}\|e\|_a.
\end{aligned}$$

Finally we get

$$\begin{aligned}
\|e\|_a^2 &\leq C\sum_{i\in\mathcal{C}} \|H\mathcal{R}(U_c)\|_{\omega_i}^2 \|\frac{1}{\sqrt{a}}\|_{L^\infty(\omega_i)}^2 \qquad (3.28)\\
&\quad + C\sum_{i\in\mathcal{F}} \left(\|\sqrt{H}\Sigma(U_{f,i})\|_{\partial\omega_i}^2 + \|h\mathcal{R}_i(U_{f,i})\|_{\omega_i}^2\right)\|\frac{1}{\sqrt{a}}\|_{L^\infty(\omega_i)}^2,
\end{aligned}$$

which proves the theorem. □

Remark 3.1 The quantity $\Sigma(U_{f,i})$ defined by (3.5) may be interpreted as a variational approximation of $n\cdot a\nabla U_{f,i}$. See [11] for further discussion on variational approximation of the flux. By defining $U_{f,i}=0$ on $\Omega\setminus\omega_i$ we extend $U_{f,i}$ to Ω. Then the jump residual $\|\sqrt{H}[\Sigma(U_{f,i})]\|_{\partial\omega_i}$ simplifies to $\|\sqrt{H}\Sigma(U_{f,i})\|_{\partial\omega_i}$. Thus the latter quantity is actually a residual quantity.

To compute $\Sigma(U_{f,i})$ we note that $\Sigma(U_{f,i})$ is a piecewise polynomial defined on the boundary of patch ω_i. Recalling that

$$(\varphi_i R(U_c), v_f)_{\omega_i} - (a\nabla U_{f,i}, \nabla v_f)_{\omega_i} = 0, \quad \text{for all } v_f \in \mathcal{V}_f^h(\omega_i), \tag{3.29}$$

we conclude that we have the same number of unknowns and equations and in practice calculating $\Sigma(U_{f,i})$ will come down to solving a linear system with a mass matrix defined on the boundary of the patch.

Remark 3.2 Since we need to calculate residuals in order to estimate the error we need to assume that discontinuities on the coefficient a are aligned with the fine scale mesh. Otherwise $\nabla \cdot a\nabla U$ can not be calculated as it stands in the estimate. If the discontinuities are not aligned we will use an approximation of a in the domains that is continuous and consider the error committed as a modelling error. More on model errors and how they can be included in error estimates see [6].

4 Special Case: Periodic Coefficients

Many multiscale applications features periodic fine scale structure. In this special case we can get more information out of our calculations. We assume that we have local scale of size ϵ and a global scale of size 1. Further we assume $a = a(x/\epsilon)$ to be smooth. If we discretize Poisson's equation with a mesh parameter $H > \epsilon$ using the standard Galerkin finite element method we have the following estimate from [8].

Proposition 4.1 *Given $0 < a < \beta$ such that $a = a(x/\epsilon)$ and $f \in L^2(\Omega)$ it holds,*

$$\|e\|_a \leq C\frac{H}{\epsilon}\|f\|. \tag{4.1}$$

Here $f \in L^2(\Omega)$. From this estimate it is clear that we can not hope to get a good approximation without resolving the fine scales. We make a similar calculation for the variational multiscale approach presented in this paper. Since the data is periodic we expect the solution to be equally hard to solve in all parts of the domain and therefore we use same sized patches with the same resolution for all patches. The fine scale mesh size is chosen to be smaller then the oscillations, $h < \epsilon$.

Theorem 4.1 *Assuming $f \in L^2(\Omega)$, $0 < a < \beta$ such that $a = a(x/\epsilon)$, and*

$u \in H^2(\Omega)$ *we have the following estimate of the error,*

$$\|e\|_a^2 \leq C \left(\frac{h}{\epsilon}\right)^2 \|f\|^2 + C \sum_{K \in \mathcal{K}} \|\sqrt{H}\Sigma(U_{f,i})\|_{\partial\omega_i}^2 \|\frac{1}{\sqrt{a}}\|_{L^\infty(\omega_i)}^2. \tag{4.2}$$

Proof. We use a global Scott-Zhang interpolant of v on the fine mesh associated with h, πv, in the following calculation,

$$\begin{aligned}
\|e\|_a^2 &= a(e,e) && (4.3)\\
&= a(e, e - \pi e) + a(e, \pi(e - \pi_c e)) && (4.4)\\
&= a(e, u - \pi u) + \sum_{i \in \mathcal{N}} a(e_c, \varphi_i \pi_{f,i}(e - e_c)) + a(e_{f,i}, \pi_{f,i}(e - \pi_c e)) && (4.5)\\
&= a(e, u - \pi u) + \sum_{i \in \mathcal{N}} (\Sigma(U_{f,i}), \pi_{f,i}(e - \pi_c e))_{\partial\omega_i} && (4.6)\\
&= \mathrm{I} + \mathrm{II}. && (4.7)
\end{aligned}$$

We start with the first part. We shall use the following result from approximation theory,

$$\|u - \pi u\|_a^2 \leq \beta^2 \|\nabla(u - \pi u)\|^2 \leq \beta^2 h^2 |u|_2^2, \tag{4.8}$$

where $|u|_2$ is the $H^2(\Omega)$ semi-norm of u and the following regularity estimate, see [8],

$$|u|_2 \leq \frac{C}{\epsilon}\|f\|. \tag{4.9}$$

Together equations (4.8) and (4.9) give,

$$\|u - \pi u\|_a \leq C\frac{h}{\epsilon}\|f\|. \tag{4.10}$$

We can now proceed with the following calculation,

$$\mathrm{I} \leq \frac{1}{4}\|e\|_a^2 + \|u - \pi u\|_a^2 \leq \frac{1}{4}\|e\|_a^2 + C\left(\frac{h}{\epsilon}\right)^2 \|f\|^2. \tag{4.11}$$

For the second part, II in equation (4.3), we use Cauchy-Schwartz for sums, the H^1-stability of π, see [5, 4], and $u_f = u - u_c$ to get,

$$\mathrm{II} \leq \sum_{i \in \mathcal{N}} \sqrt{H}\|\Sigma(U_{f,i})\|_{\partial\omega_i} \|\frac{1}{\sqrt{H}}\pi_{f,i}(e - \pi_c e)\|_{\omega_i}. \tag{4.12}$$

Using Cauchy-Schwartz for sums, equations (3.20), and (3.24) we get,

$$\text{II} \leq C \left(\sum_{i \in \mathcal{N}} \|\sqrt{H}\Sigma(U_{f,i})\|^2 \|\frac{1}{\sqrt{a}}\|^2_{L^\infty(\omega_i)} \right)^{1/2} \|e\|_a \tag{4.13}$$

$$\leq C \sum_{i \in \mathcal{N}} \|\sqrt{H}\Sigma(U_{f,i})\|^2 \|\frac{1}{\sqrt{a}}\|^2_{L^\infty(\omega_i)} + \frac{1}{4}\|e\|^2_a. \tag{4.14}$$

Finally, combining equations (4.7), (4.11), and (4.14) the theorem follows. □

Again we see that the error committed by the restriction to patches is measured by the size of the normal derivative of the fine scale solution on the boundary of the patch. This quantity will go to zero when the patch size equals the entire domain which means that the method is guaranteed to converge to the reference solution on the global fine mesh.

When solving problems with periodic coefficients it is very reasonable to choose structured meshes if possible aligned with the oscillations. It is also natural to put in equal computational effort in all parts of the domain which means that no adaptivity in terms of where the fine scale should be solved is necessary. However, it is still important to choose the relation between the parameters h and L adaptively.

If the mesh is aligned with the oscillations local problems solved inside the domain will differ only in the right hand side which means that the computational effort is almost negligible. If f is also periodic we would get identical contributions to the modifying matrix T and vector d in equation (2.10) from the local calculations.

Patches including parts of the boundary will also appear repeatedly in the calculations, one of each kind needs to be computed.

5 Adaptive Algorithm

We shall now use use the energy norm estimate in Theorem 3.1 to construct an adaptive algorithm. Recalling the estimate

$$\|e\|^2_a \leq C \sum_{i \in \mathcal{C}} \|H\mathcal{R}(U_c)\|^2_{\omega_i} \|\frac{1}{\sqrt{a}}\|^2_{L^\infty(\omega_i)} \tag{5.1}$$

$$+ C \sum_{i \in \mathcal{F}} \left(\|\sqrt{H}\Sigma(U_{f,i})\|^2_{\partial\omega_i} + \|h\mathcal{R}_i(U_{f,i})\|^2_{\omega_i} \right) \|\frac{1}{\sqrt{a}}\|^2_{L^\infty(\omega_i)}.$$

The three contributions to the error can easily be understood. The first term is the standard a posteriori error estimate for a Galerkin solution on the coarse mesh i.e. this is what we get if we do not solve any local problems.

The first part of the second sum represents the error arising from the fact that we solve the local problems on patches ω_i instead of the entire domain. Remember that $\Sigma(U_{f,i})$ is closely related to the normal derivative of the fine scale solution on the boundary of the patches. Finally, the second part of the second sum represents the fine scale resolution. The two contributions to the second sum clearly points out the parameters of interest when using our method. The first one is the patch size, increasing patch size will decrease $\|\sqrt{H}\Sigma_i(U_{f,i})\|_{\partial\omega_i}$ since $\Sigma_i(U_{f,i})$ is an approximation of the normal derivative of $U_{f,i}$ which decreases as the patch size increases due to the localized load in the right hand side of the fine scale equations (2.13) and the split of spaces which forces $U_{f,i}$ to be zero in all coarse nodes. The second part represents the effect of discretizing the local problems using the fine scale mesh size h.

Based on the error estimate (5.1) we now state the following adaptive algorithm:

Adaptive Algorithm.

1. Start with no nodes in $\mathcal{F}$.
2. Calculate a solution U_c on the coarse mesh by solving (2.13).
3. For all $i \in \mathcal{C}$ calculate the residuals for each coarse node, $R_i = \|H\mathcal{R}(U_c)\|^2_{\omega_i}$.
4. For all $i \in \mathcal{F}$ calculate the contributions from the first term of the local problems, $S_i = \|\sqrt{H}\Sigma(U_{f,i})\|^2_{\partial\omega_i}$ and the second term, $W_i = \|h\mathcal{R}_i(U_{f,i})\|^2_{\omega_i}$.
5. For large values in R_i add i to $\mathcal{F}$, for large values in S_i or W_i either increase the number of layers or decrease the fine scale mesh size h for local problem i. Return to 2 or stop if the desired tolerance is reached.

The idea is to start with an adaptively refined mesh on the coarse scale that is derived using standard techniques with refinement and coarsening. When we finally start to use the algorithm above and solve local problems it is also possible to use coarsening in areas where the solution is too good. This can either be done by increasing the local mesh parameter h or decreasing the patch size L and it is done according to the indicators S_i and W_i.

6 Numerical Examples

We solve two dimensional model problems with piecewise linear approximation defined on a uniform triangular mesh.

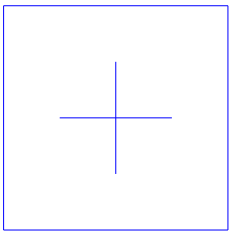


Figure 3: Unit square with slits between $(0.25, 0.5)$ and $(0.75, 0.5)$ and between $(0.5, 0.25)$ and $(0.5, 0.75)$.

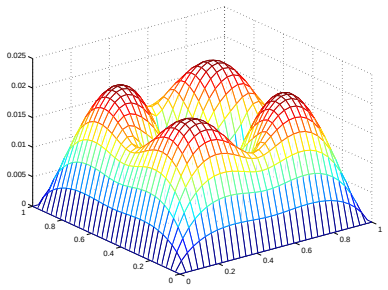


Figure 4: The solution calculated using 1089 nodes.

Example 1. In the first example we let $a = 1$, $f = 1$, and Ω be the unit square with a crack that forms a plus sign, see Figure 3. The solution u is forced to be zero on the boundary including the slits, see Figure 4. We solve the problem by using the adaptive algorithm above with a refinement level of about 20 % in each iteration. We start with one refinement and one layer stars for the local problems. Figure 5 (left) shown the adaptive choice of coarse nodes for which local problems needs to be solved.

After the second iteration no more local problems are added but the number of layers is increased to two, see Figure 5 (right). As seen the algorithm decides to increase the number of layers for all coarse nodes. This indicates that the normal derivative of $U_{f,i}$ is not small on the boundary of the patches. This is exactly what we get if we study a specific choice of $U_{f,i}$ with center close to the cracks, see Figure 6. These solutions looks quite different from the ones that origins from a smooth region found in Figure 2. The local contribution has a constant sign which indicates a constant signed error in the Galerkin solution. In Figure 7 we study the error compared to a reference solution of the standard Galerkin solution and the solutions after one and two iterations. We use the same scale in all three plots. The

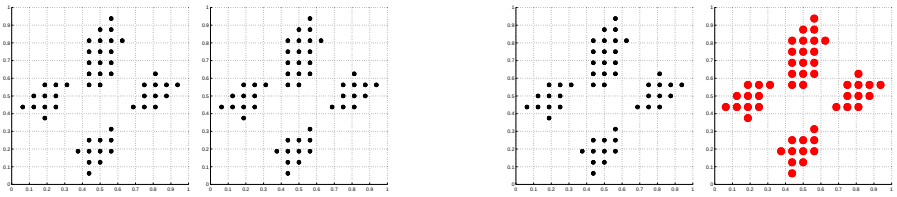


Figure 5: Refinement level, $h = H \cdot 2^{-k}$ (left), and number of layers L (right) for each coarse node after the first and second iteration. We have $k = 1$ after both iterations and $L = 1$ after the first and $L = 2$ after the second.

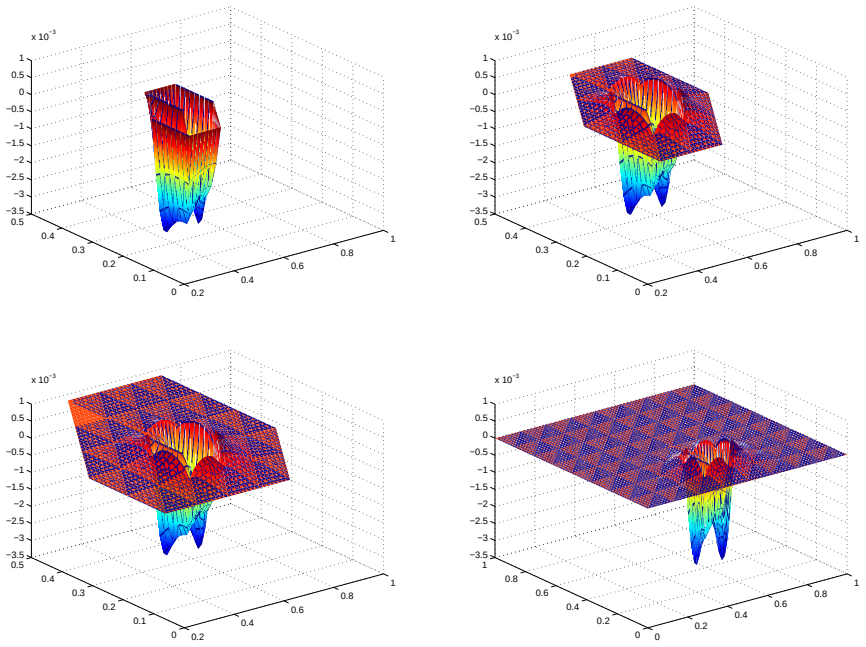


Figure 6: Localized solution $U_{f,i}$ to the fine scale equations in a rough region using one, two, three layer stars, and the entire domain.

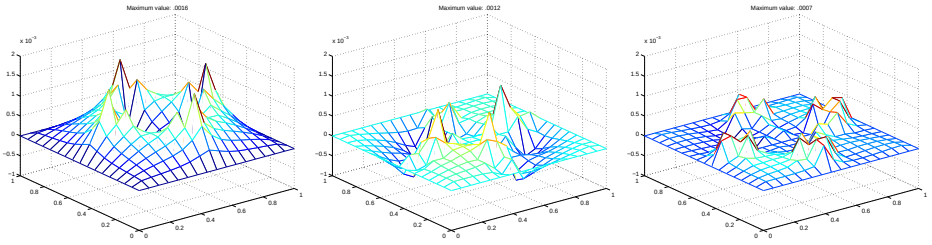


Figure 7: The error in the Galerkin solution (left), after one step in the adaptive algorithm (middle), and after two steps (right).

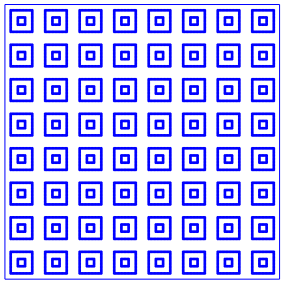


Figure 8: The coefficient is discontinuous with the values $a = 1$ on the white areas and $a = 0.05$ on the dark areas.

Galerkin solution has large error in the singularities and the error is positive. We see that the local problems decreases the error in each iteration.

Example 2. In this example we use a simple geometry, the unit square, but we let the coefficient a oscillate rapidly with period $\epsilon = 1/8$ according to Figure 8. The thickness of the black squares with $a = 0.05$ in Figure 8 is $1/96$ and this is the scale we need to resolve in order to get a good approximation. We calculate a reference solution on the fine space there $h = 1/192$ and compare it to standard Galerkin calculated using quadrature on the coarse mesh $H = 1/48$ with and without solving local problems. We see that standard Galerkin on a coarse mesh performs badly for this problem, see Figure 9. This is not surprising. From equation (4.1) we see that we have no control of the error what so ever when using Galerkin if the fine scale features are not resolved. In this case $H = 1/48$ but the fine scale features are of size $1/96$ ever though the period of repetition is $1/8$.

The approximation calculated using two layer stars on the other hand

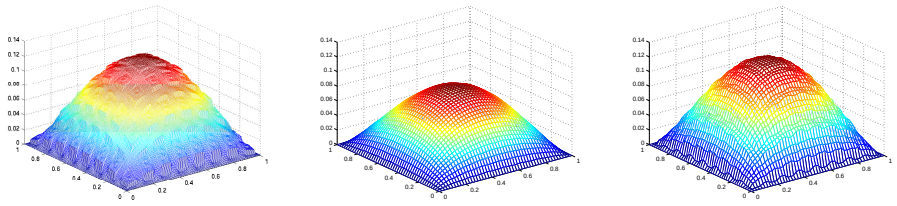


Figure 9: Reference solution (left), standard Galerkin on coarse mesh (middle), and solution with local problems using two layer stars (right).

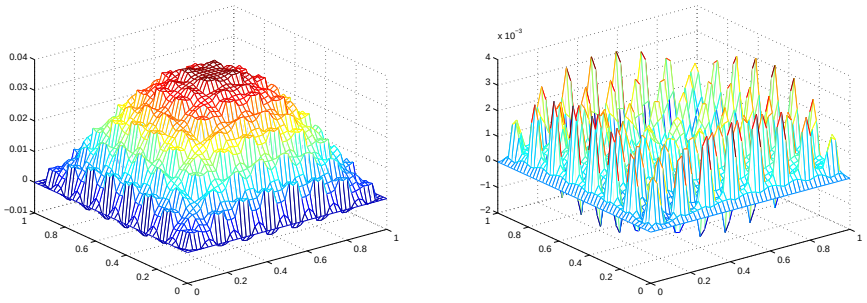


Figure 10: Error for standard Galerkin (left) and for the solution using local problems with two layer stars (right).

gives a very nice improvement of the solution. Here the fine scale features are resolved and the error committed are only due to the restriction to patches. The magnitude is correct and if we study the error between this solution and the reference solution we get a ten times smaller error, see Figure 10. The coarse mesh is aligned to the oscillations in a so this nice solution can be calculated by solving extremely few small localized problems at a very low cost.

As mentioned before calculating a modified stiffness matrix rather than using an iterative approach is very efficient in the periodic setting. To understand the method it is interesting to know how the method actually modifies the stiffness matrix. We do this by studying the spectrum of the resulting matrix $A + T$ for different number of layers in Figure 11. We study the twenty lowest and most significant eigenvalues. The first thing we note is that the eigenvalues of $A + T$ always are smaller than the ones of A. This is natural since the discretization increases eigenvalues of the operator. We also see that already after two layers we get very nice agreement with the

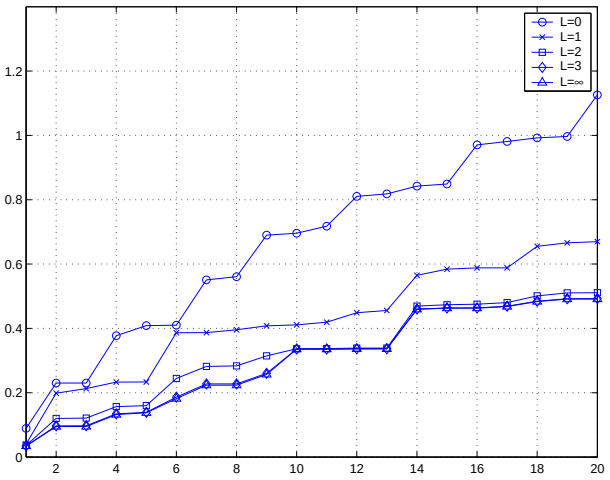


Figure 11: The twenty lowest eigenvalues of the matrix $A + T$ for fine scale problems solved using no stars, one layer stars, two layer stars, three layer stars, and on the entire domain.

correct spectrum we like to approximate.

From our numerical examples we conclude that using one layer stars almost never is enough to get good accuracy but already two layer stars gives very nice improvement of the standard Galerkin solution.

References

[1] T. Arbogast and S. L. Bryant, A two-scale numerical subgrid technique for waterflood simulations, SPE J., pp 446-457, Dec. 2002.

[2] I. Babuška and A. Miller, *A feedback finite element method with a posteriori error estimations: Part I. The finite element method and some basic properties of the a posteriori error estimator,* Comp. Meth. Appl. Mech. Eng., 61, 1-40, 1987.

[3] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods,* Springer Verlag, 1994.

[4] C. Carstensen, *Merging the Bramble-Pasciak-Steinbach and the Crouzeix-Thomée criterion for the H^1-stability of the L^2-projection onto finite element spaces,* Math. Comp., 71, 237, 157-163, 2002.

[5] M. Crouzeix and V. Thomée, *The Stability in L_p and W_p^1 of the L_2-projection onto Finite Element Function Spaces,* Mathematics of Computation, 48, 178, 521-532, 1987.

[6] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, *Computational differential equations,* Studentlitteratur, 1996.

[7] T. Y. Hou and X.-H. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media,* J. Comput. Phys. 134, 169-189, 1997.

[8] T. Y. Hou and X.-H. Wu, and Z. Cai, *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients,* Math. Comp., 68, 227, 913-943, 1999.

[9] T. J. R. Hughes, *Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods,* Comput. Methods Appl. Mech. Engrg. 127, 387-401, 1995.

[10] T. J. R. Hughes, G. R. Feijóo, L. Mazzei and J.-B. Quincy, *The variational multiscale method - a paradigm for computational mechanics,* Comput. Methods Appl. Mech. Engrg. 166, 3-24, 1998.

[11] T. J. R. Hughes, L Mazzei, A. A. Oberai, and M. G. Larson, *The continuous Galerkin method is locally conservative,* J. Comput. Phys., Vol. 163(2), 467-488, 2000.

[12] M. G. Larson and A. Målqvist, *Adaptive variational multiscale methods based on a posteriori error estimation: Duality techniques for elliptic problems,* Lecture notes in: Computational Science and Engineering. B. Enquist, P. Löstedt, O. Runborg (eds.) Vol. 44, 181-193, 2005.

[13] M. G. Larson and A. Målqvist, *Adaptive variational multiscale methods based on a posteriori error estimation,* Proceedings of the 4th European Congress on Computational Methods in Applied Sciences and Engineering, P. Neittaanmäki, T. Rossi, S. Korotov, E. Onate, J. Periaux, and D. Knörzer (eds.), 2004.

[14] P. Morin, R. H. Nochetto, and K. G. Siebert, *Local problems on stars: A posteriori error estimators, convergence, and performance,* Math. Comp., 72, 1067-1097, 2003.

[15] A. A. Oberai and P. M. Pinsky, *A multiscale finite element method for the Helmholtz equation,* Comput. Methods Appl. Mech. Engrg. 154, 281-297, 1998.

Adaptive Variational Multiscale Methods Based on A Posteriori Error Estimation: Duality Techniques for Elliptic Problems

Mats G. Larson ** and Axel Målqvist

Department of Mathematics
Chalmers University of Technology
S-412 96, Göteborg, Sweden
mgl@math.chalmers.se, axel@math.chalmers.se

Summary. The variational multiscale method (VMS) provides a general framework for construction of multiscale finite element methods. In this paper we propose a method for parallel solution of the fine scale problem based on localized Dirichlet problems which are solved numerically. Next we present a posteriori error representation formulas for VMS which relates the error in linear functionals to the coarse scale discretization error, resolution of fine scales, and size of the patches in the localized fine scale problems. These formulas are derived by using duality techniques. Based on the a posteriori error representation formula we propose an adaptive VMS with automatic tuning of the critical parameters. We primarily study elliptic second order partial differential equations with highly oscillating coefficients or localized singularities.

1 Introduction

Many problems in science and engineering involve models of physical systems on many scales. For instance, models of materials with microstructure such as composites and flow in porous media. In such problems it is in general not feasible to seek for a numerical solution which resolves all scales. Instead we may seek to develop algorithms based on a suitable combination with a global problem capturing the main features of the solution and localized problems which resolves the fine scales. Since the fine scale problems are localized the computation on the fine scales is parallel in nature.

Previous work.

The Variational Multiscale Method is a general framework for derivation of basic multiscale method in a variational context, see Hughes [7, 8]. The basic

**Corresponding author

idea is to decompose the solution into fine and coarse scale contributions, solve the fine scale equation in terms of the residual of the coarse scale solution, and finally eliminate the fine scale solution from the coarse scale equation. This procedure leads to a modified coarse scale equation where the modification accounts for the effect of fine scale behavior on the coarse scales. In practice it is necessary to approximate the fine scale equation to make the method realistic. In several works various ways of analytical modeling are investigated often based on bubbles or element Green's functions, see Oberai and Pinsky, [11] and Arbogast [4]. In [6] Hou and Wu present a different approach. Here the fine scale equations are solved numerically on a finer mesh. The fine scale equations are solved inside coarse elements and are thus totally decoupled.

New contributions.

In this work we present a simple technique for numerical approximation of the fine scale equation in the variational multiscale method. The basic idea is to split the fine scale residual into localized contributions using a partition of unity and solving corresponding decoupled localized problems on patches with homogeneous Dirichlet boundary conditions. The fine scale solution is approximated by the sum $U_f = \sum_i U_{f,i}$ of the solutions $U_{f,i}$ to the localized problems. The accuracy of U_f depends on the fine scale mesh size h and the size of the patches. We note that the fine scale computation is naturally parallel.

To optimize performance we seek to construct an adaptive algorithm for automatic control of the coarse mesh size H, the fine mesh size h, and the size of the patches. Our algorithm is based on the following a posteriori estimate of the error $e = u - U_c - U_f$ for the Poisson equation with variable coefficient a:

$$(e,\psi) = \sum_{i\in\mathcal{C}} (\varphi_i R(U_c), \phi_f) + \sum_{i\in\mathcal{F}} \left((\varphi_i R(U_c), \phi_f) - a(U_{f,i}, \phi_f) \right), \qquad (1)$$

where $\psi \in H^{-1}(\Omega)$ is a given distribution, $\mathcal{C}$ refers to nodes where no local problems have been solved, $\mathcal{F}$ to nodes where local problems are solved, U_c is the coarse scale solution, $U = U_c + U_f$, $R(U) = f + \nabla \cdot a\nabla U$ is the residual, $\{\varphi_i\}_{i\in\mathcal{C}\cup\mathcal{F}}$ are coarse base functions, and ϕ_f is the fine scale part of a dual solution driven by ψ.

If no fine scale equations are solved we only obtain the first term in the estimate. The second term relates the fine scale mesh parameter h to the patch size ω_i on which the local problems are solved. We have derived a similar estimate for the error in energy norm, see [10].

The framework is fairly general and may be extended to other types of multiscale methods, for instance, based on localized Neumann problems.

Outline.

First we introduce the model problem and the variational multiscale formulation of this problem, we also discuss the split of the coarse and fine scale

spaces. In the following section we present a posteriori estimates of the error. These results leads to an adaptive algorithm. We present numerical results and finally we present concluding remarks and suggestions on future work.

2 The Variational Multiscale Method

2.1 Model Problem

We study the Poisson equation with a highly oscillating coefficient a and homogeneous Dirichlet boundary conditions: find $u \in H_0^1(\Omega)$ such that

$$-\nabla \cdot a\nabla u = f \quad \text{in } \Omega, \tag{2}$$

where Ω is a polygonal domain in $\mathbf{R}^d$, $d = 1, 2$, or 3 with boundary Γ, $f \in L^2(\Omega)$, and $a \in L^\infty(\Omega)$ such that $a(x) \geq \alpha_0 > 0$ for all $x \in \Omega$. The variational form of (2) reads: find $u \in \mathcal{V} = H_0^1(\Omega)$ such that

$$a(u, v) = (f, v) \quad \text{for all } v \in \mathcal{V}, \tag{3}$$

with the bilinear form

$$a(u, v) = (a\nabla u, \nabla v), \tag{4}$$

for all $u, v \in \mathcal{V}$.

2.2 The Variational Multiscale Method

We employ the variational multiscale scale formulation, proposed by Hughes see [7, 8] for an overview, and introduce a coarse and a fine scale in the problem. We choose two spaces $\mathcal{V}_c \subset \mathcal{V}$ and $\mathcal{V}_f \subset \mathcal{V}$ such that

$$\mathcal{V} = \mathcal{V}_c \oplus \mathcal{V}_f. \tag{5}$$

Then we may pose (3) in the following way: find $u_c \in \mathcal{V}_c$ and $u_f \in \mathcal{V}_f$ such that

$$\begin{aligned} a(u_c, v_c) + a(u_f, v_c) &= (f, v_c) \quad \text{for all } v_c \in \mathcal{V}_c, \\ a(u_c, v_f) + a(u_f, v_f) &= (f, v_f) \quad \text{for all } v_f \in \mathcal{V}_f. \end{aligned} \tag{6}$$

Introducing the residual $R : \mathcal{V} \to \mathcal{V}'$ defined by

$$(R(v), w) = (f, w) - a(v, w) \quad \text{for all } w \in \mathcal{V}, \tag{7}$$

the fine scale equation takes the form: find $u_f \in \mathcal{V}_f$ such that

$$a(u_f, v_f) = (R(u_c), v_f) \quad \text{for all } v_f \in \mathcal{V}_f. \tag{8}$$

Thus the fine scale solution is driven by the residual of the coarse scale solution. Denoting the solution u_f to (8) by $u_f = TR(u_c)$ we get the modified coarse scale problem

$$a(u_c, v_c) + a(TR(u_c), v_c) = (f, v_c) \quad \text{for all } v_c \in \mathcal{V}_c. \tag{9}$$

Here the second term on the left hand side accounts for the effects of fine scales on the coarse scales.

2.3 A VMS Based on Localized Dirichlet Problems

We introduce a partition $\mathcal{K} = \{K\}$ of the domain Ω into shape regular elements K of diameter H_K and we let $\mathcal{N}$ be the set of nodes. Further we let $\mathcal{V}_c$ be the space of continuous piecewise polynomials of degree p defined on $\mathcal{K}$.

We shall now construct an algorithm which approximates the fine scale equation by a set of decoupled localized problems. We begin by writing $u_f = \sum_{i\in\mathcal{N}} u_{f,i}$ where

$$a(u_{f,i}, v_f) = (\varphi_i R(u_c), v_f) \quad \text{for all } v_f \in \mathcal{V}_f, \tag{10}$$

and $\{\varphi_i\}_{i\in\mathcal{N}}$ is the set of Lagrange basis functions in $\mathcal{V}_c$. Note that $\{\varphi_i\}_{i\in\mathcal{N}}$ is a partition of unity with support on the elements sharing the node i. We call the set of elements with one corner in node i a mesh star in node i and denote it S_1^i. Thus functions $u_{f,i}$ correspond to the fine scale effects created by the localized residuals $\varphi_i R(u_c)$. Introduce this expansion of u_f in the right hand side of the fine scale equation (6) and get: find $u_c \in \mathcal{V}_c$ and $u_f = \sum_{i\in\mathcal{N}} u_{f,i} \in \mathcal{V}_f$ such that

$$\begin{aligned} a(u_c, v_c) + a(u_f, v_c) &= (f, v_c) \quad \text{for all } v_c \in \mathcal{V}_c, \\ a(u_{f,i}, v_f) &= (\varphi_i R(u_c), v_f) \quad \text{for all } v_f \in \mathcal{V}_f \text{ and } i \in \mathcal{N}. \end{aligned} \tag{11}$$

We use this fact to construct a finite element method for solving (11) approximately in two steps.

- For each coarse node we define a patch ω_i such that $\text{supp}(\varphi_i) \subset \omega_i \subset \Omega$. We denote the boundary of ω_i by $\partial\omega_i$.
- On these patches we define piecewise polynomial spaces $\mathcal{V}_f^h(\omega_i)$ with respect to a fine mesh with mesh function $h = h(x)$ defined as a piecewise constant function on the fine mesh. Functions in $\mathcal{V}_f^h(\omega_i)$ are equal to zero on $\partial\omega_i$.

The resulting method reads: find $U_c \in \mathcal{V}_c$ and $U_f = \sum_i^n U_{f,i}$ where $U_{f,i} \in \mathcal{V}_f^h(\omega_i)$ such that

$$\begin{aligned} a(U_c, v_c) + a(U_f, v_c) &= (f, v_c) \quad \text{for all } v_c \in \mathcal{V}_c, \\ a(U_{f,i}, v_f) &= (\varphi_i R(U_c), v_f) \quad \text{for all } v_f \in \mathcal{V}_f^h(\omega_i) \text{ and } i \in \mathcal{N}. \end{aligned} \tag{12}$$

Since the functions in the local finite element spaces $\mathcal{V}_f^h(\omega_i)$ are equal to zero on $\partial\omega_i$, U_f and therefore $U = U_c + U_f$ will be continuous.

Remark 2.1 *For problems with multiscale phenomena on a part of the domain it is not necessary to solve local problems for all coarse nodes. We let $\mathcal{C} \subset \mathcal{N}$ refer to nodes where no local problems are solved and $\mathcal{F} \subset \mathcal{N}$ refer to nodes where local problems are solved. Obviously $\mathcal{C} \cup \mathcal{F} = \mathcal{N}$. We let $U_{f,i} = 0$ for $i \in \mathcal{C}$.*

Remark 2.2 *The choice of the subdomains ω_i is crucial for the method. We introduce a notation for extended stars of many layers of coarse elements recursively in the following way. Let S_1^i be the support of the coarse scale Lagrangian base function φ_i in node i. The extended mesh star $S_L^i = \cup_{x_j \in S_{L-1}^i} S_1^j$ for $L > 1$, where $\varphi_j(x_j) = 1$, nodes on the boundary are included. We refer to L as the numbers of layers, see Figure 1.*

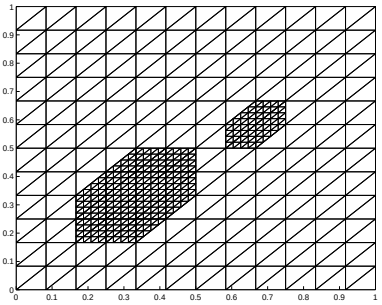


Fig. 1. Two (left) and one (right) layer stars.

In Figure (2) we plot solutions to localized fine scale problems $U_{f,i}$ on different patches. We note how $U_{f,i}$ decays rapidly outside the support of φ_i. It appears to be enough to use two layers in this example to capture the true behavior of the fine scale solution.

2.4 Subspaces

The choice of the fine scale space $\mathcal{V}_f$ can be done in different ways. In a paper by Aksoylu and Holst [3] three suggestions are made.

Hierarchical basis method.

The first and perhaps easiest approach is to let $\mathcal{V}_f = \{v \in \mathcal{V} : v(x_j) = 0, j = \mathcal{N}\}$, where $\{x_i\}_{i \in \mathcal{N}}$ are the coarse mesh nodes. When $\mathcal{V}_f$ is discretized by the standard piecewise polynomials on the fine mesh this means that the fine scale base functions will have support on fine scale stars.

BPX preconditioner.

The second approach is to let $\mathcal{V}_f$ be $L^2(\Omega)$ orthogonal to $\mathcal{V}_c$. In this case we will have global support for the fine scale base functions but rapid decay outside the fine mesh stars.

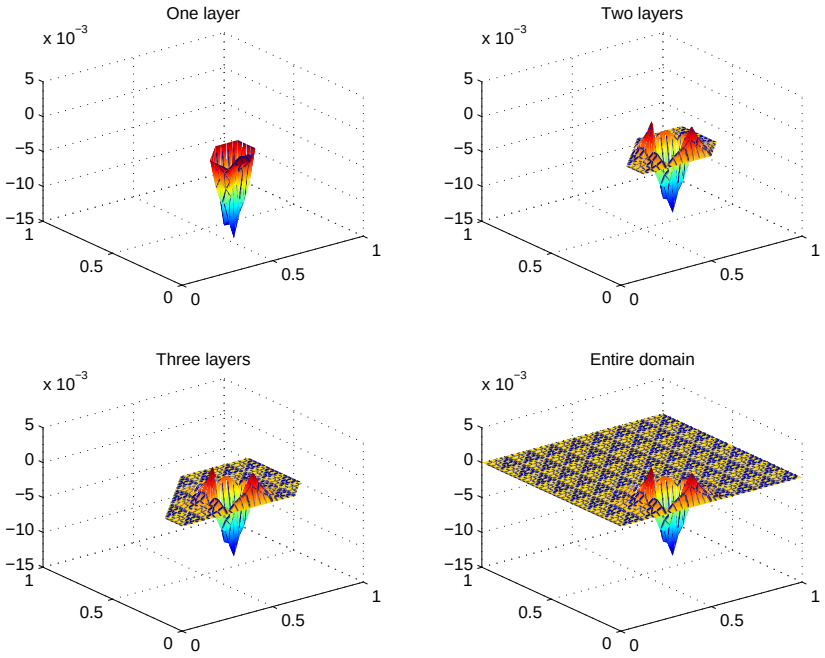


Fig. 2. The fine scale solution $U_{f,i}$ for different patches.

Wavelet modified hierarchical basis method.

The third choice is a mix of the other two. The fine scale space $\mathcal{V}_f$ is defined as an approximate $L^2(\Omega)$ orthogonal version of the Hierarchical basis method. We let $Q_c^a v \in \mathcal{V}_c$ be an approximate solution (a small number of Jacobi iterations) to

$$(Q_c^a v, w) = (v, w), \quad \text{for all } w \in \mathcal{V}_c, \tag{13}$$

and define the Wavelet modified hierarchical basis function associated with the hierarchical basis function φ_{HB} to be,

$$\varphi_{WHB} = (I - Q_c^a)\varphi_{HB}, \tag{14}$$

see Figure 3.

For an extended description of these methods see [2, 3, 1]. In this paper we focus on the WHB method.

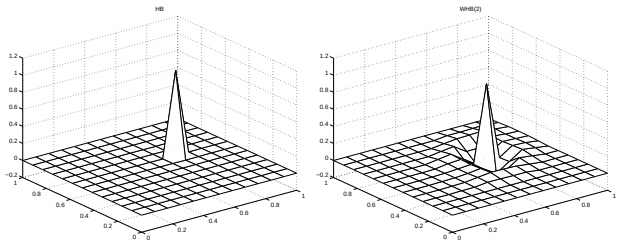


Fig. 3. HB-function and WHB-function with two Jacobi iterations.

3 A Posteriori Error Estimates

3.1 The Dual Problem

To derive a posteriori estimates of the error in a given linear functional (e, ψ), with $e = u - U$ and $\psi \in H^{-1}(\Omega)$ a given weight, we introduce the following dual problem: find $\phi \in \mathcal{V}$ such that

$$a(v, \phi) = (v, \psi) \quad \text{for all } v \in \mathcal{V}. \tag{15}$$

In the VMS setting this yields: find $\phi_c \in \mathcal{V}_c$ and $\phi_f \in \mathcal{V}_f$ such that

$$\begin{aligned} a(v_c, \phi_c) + a(v_c, \phi_f) &= (v_c, \psi), \quad \text{for all } v_c \in \mathcal{V}_c, \\ a(v_f, \phi_f) + a(v_f, \phi_c) &= (v_f, \psi), \quad \text{for all } v_f \in \mathcal{V}_f. \end{aligned} \tag{16}$$

3.2 Error Representation Formula

We now derive an error representation formula involving both the coarse scale error $e_c = u_c - U_c$ and the fine scale error $e_f = \sum_{i \in \mathcal{N}} e_{f,i} := \sum_{i \in \mathcal{N}} (u_{f,i} - U_{f,i})$ that arises from using our finite element method (12).

We use the dual problem (16) to derive an a posteriori error estimate for a linear functional of the error $e = e_c + e_f$. If we subtract the coarse part of equation (12) from the coarse part of equation (11) we get the Galerkin orthogonality,

$$a(e_c, v_c) + a(e_f, v_c) = 0 \quad \text{for all } v_c \in \mathcal{V}_c. \tag{17}$$

The same argument on the fine scale equation gives for $i \in \mathcal{F}$,

$$a(e_{f,i}, v_f) = -a(e_c, \varphi_i v_f), \quad \text{for all } v_f \in \mathcal{V}_f^h(\omega_i). \tag{18}$$

We are now ready to state an error representation formula.

Theorem 3.1 *If $\psi \in H^{-1}(\Omega)$ then,*

$$(e, \psi) = \sum_{i \in \mathcal{C}} (\varphi_i R(U_c), \phi_f) + \sum_{i \in \mathcal{F}} \left((\varphi_i R(U_c), \phi_f) - a(U_{f,i}, \phi_f) \right). \tag{19}$$

Proof. Starting from the definition of the dual problem and letting $v = e = u - U_c - U_f$ we get

$$
\begin{aligned}
(e,\psi) &= a(e,\phi) && (20)\\
&= a(e,\phi_f) && (21)\\
&= a(u-U_c,\phi_f) - a(U_f,\phi_f) && (22)\\
&= (R(U_c),\phi_f) - a(U_f,\phi_f) && (23)\\
&= (R(U_c),\phi_f) - \sum_{i\in\mathcal{F}} a(U_{f,i},\phi_f) && (24)\\
&= \sum_{i\in\mathcal{C}} (\varphi_i R(U_c),\phi_f) + \sum_{i\in\mathcal{F}} (\varphi_i R(U_c),\phi_f) - a(U_{f,i},\phi_f), && (25)
\end{aligned}
$$

which proves the theorem.

Since equation (12) holds we can subtract functions $v^h_{f,i} \in \mathcal{V}^h_f(\omega_i)$ where $i \in \mathcal{F}$ from equation (25). For example we choose $v^h_{f,i} = \pi_{h,i}\phi_f$, where $\pi_{h,i}\phi_f$ is the Scott-Zhang interpolant, see [5], of ϕ_f onto $\mathcal{V}^h_f(\omega_i)$ to get

$$
\begin{aligned}
(e,\psi) = &\sum_{i\in\mathcal{C}} (\varphi_i R(U_c),\phi_f) && (26)\\
&+ \sum_{i\in\mathcal{F}} \left((\varphi_i R(U_c),\phi_f - \pi_{h,i}\phi_f) - a(U_{f,i},\phi_f - \pi_{h,i}\phi_f) \right).
\end{aligned}
$$

Remark 3.1 *Since the dual problem defined in equation (16) is equally hard to solve as the primal problem we need to solve it numerically as well. Normally it would not be sufficient to solve the dual problem with the same accuracy as the primal due to the Galerkin Orthogonality. However in this setting things are a bit different. Calculating ϕ_f with minimum refinement (one time) on the local problems for $i \in \mathcal{N}$ will not result in an error (e,ψ) equal to zero. The important thing is to only store the part of ϕ_f with support on ω_i when calculating term i in the sum of equation (19). The entire function ϕ_f might be hard to store in the memory of the computer.*

4 Adaptive Algorithm

We use the error representation formula in Theorem 3.1 to construct an adaptive algorithm. We remember the result,

$$
(e,\psi) = \sum_{i\in\mathcal{C}} (\varphi_i R(U_c),\phi_f) + \sum_{i\in\mathcal{F}} \left((\varphi_i R(U_c),\phi_f) - a(U_{f,i},\phi_f) \right). \qquad (27)
$$

The first sum of the error representation formula is very similar to what we would get from using standard Galerkin on the coarse mesh. The function

$\phi_f = \phi - \phi_c \sim H\nabla\phi$ which is exactly what we would expect. For the second sum we have an extra orthogonality namely that from equation (26). We have $\phi_f - \pi_{h,i}\phi_f \sim h\nabla\phi$ if the patches $\omega_i = \Omega$ i.e. we get the fine scale convergence. But in practice the patches are much smaller so we end up somewhere in between h and H convergence. To sum up this discussion there are three parameters of interest that need to be considered in an adaptive algorithm, H, h, and the size of the patches.

Adaptive Algorithm.

1. Start with no nodes in $\mathcal{F}$.
2. Calculate the primal U_c by solving equation (12).
3. Calculate the dual solution locally ϕ_f with low accuracy for all coarse nodes. (ϕ_f does not need to be solved very accurately to point out the correct nodes for local calculations.)
4. Calculate the contributions to the error $C_i = (\varphi_i R(U_c), \phi_f)$ for each coarse node where no fine scale problems have been solved and the contributions $F_i = ((\varphi_i R(U_c), \phi_f) - a(U_{f,i}, \phi_f))$ to nodes where fine scale problems have been solved.
5. For large values in C_i solve more local problems and for large values in F_i either increase the number of layers or decrease the fine scale mesh size h for local problem i. Stop if the desired tolerance is reached or go to 2.

5 Numerical Examples

We solve two dimensional model problems with linear base functions defined on a uniform triangular mesh.

Example 1.

In this example we demonstrate how we can get highly improved accuracy in one part of the domain by choosing the load in the dual problem ψ equal to the indicator function for this domain. We consider the unit square with a crack in the form of a plus sign on which the solution is forced to be zero, see Figure 4 (left). We let ψ be equal to one in the lower left quadrant (marked with a thin lattice in the figure) and zero in the rest of the domain. To the right in Figure 4 we see a reference solution to the Poisson equation with $a = f = 1$ and homogeneous Dirichlet boundary conditions on this geometry. The idea is to use the adaptive algorithm to choose which areas that needs to be solved with higher accuracy. In Figure 5 we plot the dual solution, with ψ chosen as described above, to the left and the fine scale part of the dual solution to the right. After two iterations in the adaptive algorithm we see clearly that local problems have only nodes in the lower left corner. In Figure 6 the small circles refer to fine scale problems solved with two layer stars and the bigger circles refer to fine scale problems solved with three layer stars. The

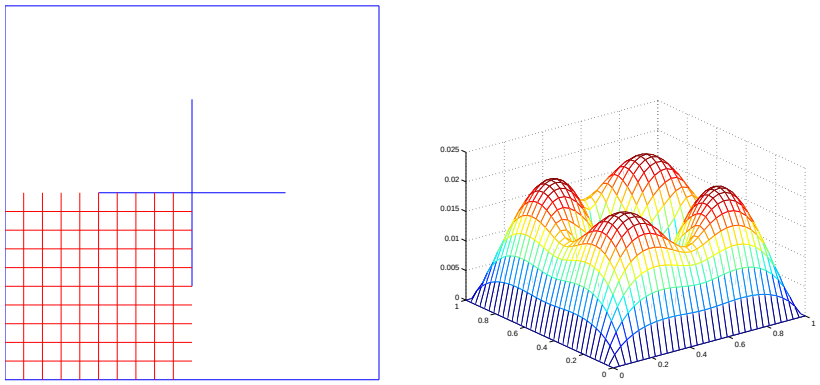


Fig. 4. Geometry (left) and Reference solution (right).

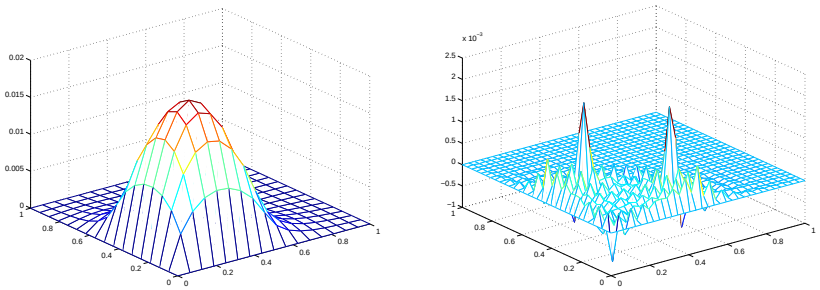


Fig. 5. Dual solution (left) and fine scale part of the dual solution (right) calculated with $\psi = I_{\{0\leq x,y\leq 0.5\}}$.

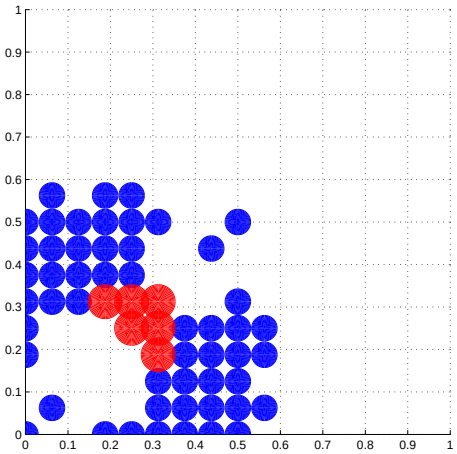


Fig. 6. Local problems solved with two and three layer stars.

improvement in the solution after two iterations in the adaptive algorithm is very clear. In Figure 7 we compare the standard Galerkin solution and the adaptive solution to a reference solution. We see how the error in the lower

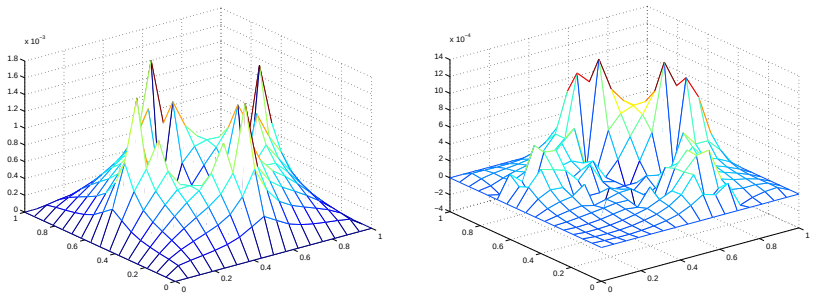


Fig. 7. Galerkin error (left) and adaptive variational multiscale method error (right).

left quadrant is much smaller but the error in the rest of the domain is very similar to the standard Galerkin error.

Example 2.

Next we turn our attention to a model problem with oscillating coefficient a in a part of the domain, see Figure 8. In this example we choose $f = \psi = 1$

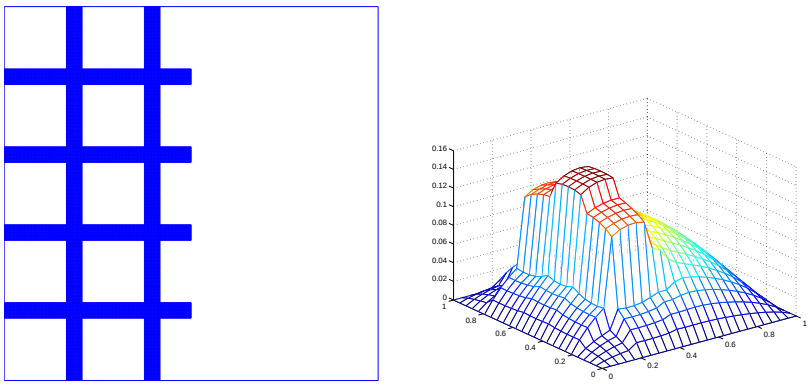


Fig. 8. The coefficient $a = 1$ on the white parts and $a = 0.05$ on the lattice (left) and reference solution on this geometry (right).

which makes the primal and the dual equivalent. In Figure 9 we note that the adaptive algorithm automatically picks nodes in the left part of the domain

for local problems to increase accuracy. In the first example we want to refine

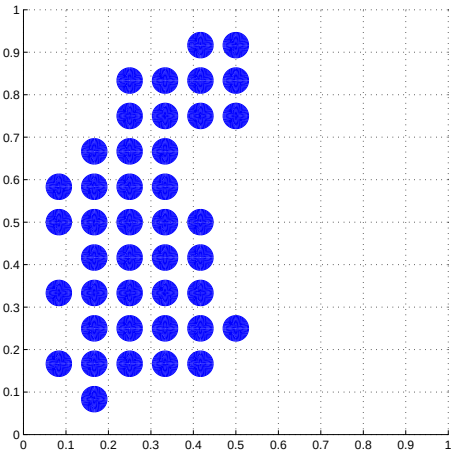


Fig. 9. The dots marks coarse nodes where local problems have been solved.

a certain part of the domain and therefore we choose ψ in order to do so, here we want good accuracy in the whole domain and the adaptive algorithm chooses where to refine automatically. Again we compare standard Galerkin and our solution to a reference solution calculated on a finer mesh. The result can be seen in Figure 10. Again we see a nice improvement compared to the

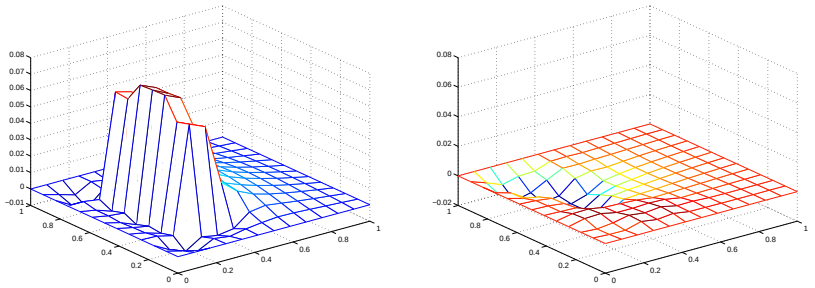


Fig. 10. Standard Galerkin error (left) and the error using adaptivity (right).

standard Galerkin error. The choice $\psi = 1$ indicates control of the mean of the error over the domain.

6 Conclusions and Future Work

We have presented a method for parallel solution of the fine scale equations in the variational multiscale method based on solution of localized Dirichlet problems on patches and developed an a posteriori error analysis for the method. Based on the estimates we design a basic adaptive algorithm for automatic tuning of the critical parameters: resolution and size of patches in the fine scale problems. It is also possible to decide whether a fine scale computation is necessary or not and thus the proposed scheme may be combined with a standard adaptive algorithm on the coarse scales. The method is thus very general in nature and may be applied to any problem where adaptivity is needed.

In this paper we have focused on two scales in two spatial dimensions. A natural extension would be to solve three dimensional problems with multiple scales. It is also natural to extend this theory to other equations modeling for instance flow and materials. We also intend to study non-linear and time dependent equations using this approach.

References

1. B. Aksoylu, S. Bond, and M. Holst *An odyssey into local refinement and multilevel preconditioning III: Implementation of numerical experiments,* SIAM J. Sci. Comput., 25(2003), 478-498.
2. B. Aksoylu and M. Holst *An odyssey into local refinement and multilevel preconditioning I: Optimality of the BPX preconditioner,* SIAM J. Numer. Anal. (2002) in review
3. B. Aksoylu and M. Holst *An odyssey into local refinement and multilevel preconditioning II: stabilizing hierarchical basis methods,* SIAM J. Numer. Anal. in review
4. T. Arbogast and S. L. Bryant, A two-scale numerical subgrid technique for waterflood simulations, SPE J., Dec. 2002, pp 446-457.
5. S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods,* Springer Verlag, 1994.
6. T. Y. Hou and X.-H. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media,* J. Comput. Phys. 134 (1997) 169-189.
7. T. J.R. Hughes, *Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods,* Comput. Methods Appl. Mech. Engrg. 127 (1995) 387-401.
8. T. J.R. Hughes, G. R. Feijóo, L. Mazzei and Jean-Baptiste Quincy, *The variational multiscale method - a paradigm for computational mechanics,* Comput. Methods Appl. Mech. Engrg. 166 (1998) 3-24.
9. T. J.R. Hughes, L Mazzei, A. A. Oberai, and M. G. Larson, *The continuous Galerkin method is locally conservative,* J. Comput. Phys., Vol. 163(2) (2000) 467-488.
10. M. G. Larson and A. Målqvist, *Adaptive Variational Multiscale Methods Based on A Posteriori Error Estimates: Energy Norm Estimates for Elliptic Problems,* submitted.

11. A. A. Oberai and P. M. Pinsky, *A multiscale finite element method for the Helmholtz equation,* Comput. Methods Appl. Mech. Engrg. 154 (1998) 281-297.

A Posteriori Error Estimates for Mixed Finite Element Approximations of Elliptic Problems

Axel Målqvist* and Mats G. Larson†

August 19, 2005

Abstract

We derive residual based a posteriori error estimates of the flux in L^2-norm for a general class of mixed methods for elliptic problems. The estimate is applicable to standard mixed methods such as the Raviart-Thomas-Nedelec and Taylor-Hood elements, as well as stabilized methods such as the Galerkin-Least squares method. The element residual in the estimate employs an elementwise computable postprocessed approximation of the displacement which gives optimal order.

1 Introduction

The Model Problem. We consider the mixed formulation of the Poisson equation with Neumann boundary conditions:

$$\begin{cases} \boldsymbol{\sigma} - \nabla u = 0 & \text{in } \Omega, \\ -\nabla \cdot \boldsymbol{\sigma} = f & \text{in } \Omega, \\ \boldsymbol{n} \cdot \boldsymbol{\sigma} = 0 & \text{on } \Gamma, \end{cases} \tag{1.1}$$

where Ω is a polygonal domain in $\mathbf{R}^n$ with boundary Γ. Assuming $\int_\Omega f\,dx = 0$, we get a well posed problem with a solution $u \in H^1(\Omega)/\mathbf{R}$ and $\boldsymbol{\sigma} \in \boldsymbol{V} = \{\boldsymbol{v} \in H(\mathrm{div};\Omega) : \boldsymbol{n} \cdot \boldsymbol{v} = 0 \text{ on } \Gamma\}$. See [9] for definitions of these function spaces.

*Graduate Research Assistant, Department of Mathematics, Chalmers University of Technology, Göteborg, axel@math.chalmers.se.

†Corresponding author, Professor of Applied Mathematics, Department of Mathematics, Umeå University, Umeå, mats.larson@math.umu.se.

Previous Work. Several works present a posteriori error estimates for mixed methods. In Carstensen [10] an error estimate in the $H(\mathrm{div};\Omega)$ norm of the flux is presented. The $H(\mathrm{div};\Omega)$ norm may be dominated by the div-part which is directly computable. When it comes to error estimates of the flux in L^2 norm of methods using richer spaces for the flux $\boldsymbol{\sigma}$ than the displacement u, such as Raviart-Thomas-Nedelec (RTN) elements, there are known difficulties. Braess and Verfürth presents a suboptimal estimate in [4]. The reason for the suboptimality is that the natural residual that arises from the first equation $\boldsymbol{\sigma} - \nabla u = 0$ in problem (1.1) may be large if the flux space is richer than the displacement space. In a recent paper Lovadina and Stenberg [13] derive an a posteriori error estimate of the L^2-norm of the flux for the RTN based methods which employs a particular postprocessed approximation of U. The proof is based on a posteriori error analysis of an equivalent method which involves the postprocessed approximation of U.

New Contributions. We derive a general a posteriori error estimate in the energy norm which is applicable to most mixed methods including the classical inf-sup stable elements, Raviart-Thomas elements, the BDM-elements and the Taylor-Hood. Our estimate is closely related to the estimate presented by Lovadina and Stenberg [13], however, our proof is more general and also reveals the fact that one can use any piecewise polynomial approximation of the displacement when computing the residual. By a small adjustment of the argument we finally, derive an estimate for the stabilized mixed method of Masud and Hughes [14]. The same technique applies to other stabilized schemes, for instance the Galerkin least squares method.

Outline. We start by presenting finite elements and the discrete version of equation (1.1) in Section 2 then we present the a posteriori error estimates in Section 3.

2 Weak Form and the Finite Element Method

Weak Formulation. We multiply the first equation in (1.1) by a test function $\boldsymbol{v} \in \boldsymbol{V}$ and integrate by parts. The second equation in (1.1) is multiplied by a test function $w \in W = L^2(\Omega)$. The weak form reads: find $\boldsymbol{\sigma} \in \boldsymbol{V}$ and $u \in W$ such that,

$$\begin{cases} (\boldsymbol{\sigma}, \boldsymbol{v}) + (u, \nabla \cdot \boldsymbol{v}) = 0 & \text{for all } \boldsymbol{v} \in \boldsymbol{V}, \\ (-\nabla \cdot \boldsymbol{\sigma}, w) = (f, w) & \text{for all } w \in W. \end{cases} \tag{2.1}$$

Our aim is to derive a posteriori error estimates of finite element approximations $\{\boldsymbol{\Sigma}, U\}$ of the exact solution $\{\boldsymbol{\sigma}, u\}$ in the energy norm $\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0$, where $\|\cdot\|_0$ denotes the $L^2(\Omega)$ norm.

The Mixed Finite Element Method. We let $\mathcal{K} = \{K\}$ be a partition of Ω into shape regular elements of diameter h_K and define the mesh function, $h(x) : \Omega \to \mathbf{R}^+$, by letting $h(x) = h_K$ for $x \in K$.

We seek an approximate solution in discrete spaces $\boldsymbol{V}_h \subset \boldsymbol{V}$ and $W_h \subset W$ defined on the partition $\mathcal{K}$. It is well known that for finite element methods based on the standard weak form (2.1) the discrete spaces must be chosen so that the inf-sup condition, see [9], is satisfied in order to guarantee a stable method. Only rather special constructions of the discrete spaces yield stable methods. In Section 3.3 we consider a stabilized mixed finite element method based on a modified weak formulation which can be based on standard continuous piecewise polynomials. We summarize some of the most well known choices of stable discrete spaces on triangles and tetrahedra for a given integer $k \geq 1$:

- *Raviart-Thomas-Nedelec* (RTN) elements, see [16, 15],
  $\boldsymbol{V}_h = \{\boldsymbol{v} \in H(\mathrm{div};\Omega) : \boldsymbol{v}|_K \in [P_{k-1}(K)]^n \oplus \boldsymbol{x}\tilde{P}_{k-1}(K) \text{ for all } K \in \mathcal{K}\}$,
  $W_h = \{w \in L^2(\Omega) : w|_K \in P_{k-1}(K) \text{ for all } K \in \mathcal{K}\}$.

- *Brezzi-Douglas-Marini* (BDM) elements, see [8, 7],
  $\boldsymbol{V}_h = \{\boldsymbol{v} \in H(\mathrm{div};\Omega) : \boldsymbol{v}|_K \in [P_k(K)]^n \text{ for all } K \in \mathcal{K}\}$,
  $W_h = \{w \in L^2(\Omega) : w|_K \in P_{k-1}(K) \text{ for all } K \in \mathcal{K}\}$.

- *Taylor-Hood* (TH), see [12],
  $\boldsymbol{V}_h = \{\boldsymbol{v} \in C(\Omega) : \boldsymbol{v}|_K \in [P_{k+1}(K)]^n \text{ for all } K \in \mathcal{K}\}$,
  $W_h = \{w \in C(\Omega) : w|_K \in P_k(K) \text{ for all } K \in \mathcal{K}\}$.

Here $C(\Omega)$ denotes the space of continuous functions on Ω, $P_k(K)$ the space of polynomials of degree k on element K, and $\tilde{P}_k(K)$ the set of homogeneous polynomials of degree k. For a more complete account of inf-sup stable spaces we refer to Brezzi-Fortin, [9]. The norms used in this paper are standard Sobolev norms following the notation, $\|\cdot\|_{s,\omega} = \|\cdot\|_{H^s(\omega)} = \|\cdot\|_{W_2^s(\omega)}$, see [1].

The mixed finite element method reads: find $\boldsymbol{\Sigma} \in \boldsymbol{V}_h$ and $U \in W_h$ such that:

$$\begin{cases} (\boldsymbol{\Sigma}, \mathbf{v}) + (U, \nabla \cdot \mathbf{v}) = 0 & \text{for all } \boldsymbol{v} \in \boldsymbol{V}_h, \\ (-\nabla \cdot \boldsymbol{\Sigma}, w) = (f, w) & \text{for all } w \in W_h. \end{cases} \tag{2.2}$$

3 A Posteriori Error Estimates

3.1 Estimate for Standard Mixed Methods

Here we present a general a posteriori error estimate in the energy norm $\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0$ involving a piecewise polynomial function Q, which may be obtained by postprocessing U. The possibility to replace U by Q is important since it

leads to a posteriori error estimates of optimal order. We are not interested in tracking the constants in the error estimates.

Theorem 3.1 *For arbitrary $Q \in \bigoplus_{K\in\mathcal{K}} P_l(K)$, with $l \geq 0$ and $f \in L^2(\Omega)$ it holds,*

$$\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 \leq C \sum_{K\in\mathcal{K}} \left(h_K^2 \|f + \nabla\cdot\boldsymbol{\Sigma}\|_{0,K}^2 + \|\boldsymbol{\Sigma} - \nabla Q\|_{0,K}^2 + h_K^{-1} \| [Q] \|_{0,\partial K}^2 \right), \tag{3.1}$$

where the jump denoted $[\cdot]$ is the difference in function value over a face in the mesh.

Proof. Starting with the left hand side we have

$$\begin{aligned}
\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 &= (\boldsymbol{\sigma} - \boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma}) && (3.2)\\
&= (\boldsymbol{\sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma}) - (\boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma}) && (3.3)\\
&= -(u, \nabla\cdot(\boldsymbol{\sigma} - \boldsymbol{\Sigma})) - (\boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma}) && (3.4)\\
&= -(u - Q, \nabla\cdot(\boldsymbol{\sigma} - \boldsymbol{\Sigma})) + (Q, -\nabla\cdot(\boldsymbol{\sigma} - \boldsymbol{\Sigma})) - (\boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma}) && (3.5)\\
&= (u - Q, f + \nabla\cdot\boldsymbol{\Sigma}) && (3.6)\\
&\qquad + \sum_{K\in\mathcal{K}} \Big((Q, -\nabla\cdot(\boldsymbol{\sigma} - \boldsymbol{\Sigma}))_K - (\boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma})_K \Big) \\
&= I + II. && (3.7)
\end{aligned}$$

We treat the two terms in equation (3.7) separately, beginning with I. From the second part of equation (2.2) we have the Galerkin orthogonality property $(f + \nabla\cdot\boldsymbol{\Sigma}, w) = 0$ for all $w \in W_h$. We let π_h denote the Scott-Zhang interploant, see [6], onto W_h and proceed with the estimates as follows

$$\begin{aligned}
I &\leq |(f + \nabla\cdot\boldsymbol{\Sigma}, u - Q)| && (3.8)\\
&\leq \|h(f + \nabla\cdot\boldsymbol{\Sigma})\|_0 \|h^{-1}(u - Q - \pi_h(u - Q))\|_0 && (3.9)\\
&\leq C\|h(f + \nabla\cdot\boldsymbol{\Sigma})\|_0 \|\nabla(u - Q)\|_0 && (3.10)\\
&= C\|h(f + \nabla\cdot\boldsymbol{\Sigma})\|_0 \|\boldsymbol{\sigma} - \boldsymbol{\Sigma} + \boldsymbol{\Sigma} - \nabla Q\|_0 && (3.11)\\
&\leq \frac{3C^2}{2}\|h(f + \nabla\cdot\boldsymbol{\Sigma})\|_0^2 + \frac{1}{4}\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 + \frac{1}{2}\|\boldsymbol{\Sigma} - \nabla Q\|_0^2. && (3.12)
\end{aligned}$$

We now turn to the second term II in equation (3.7) and start with integration by parts,

$$II = \sum_{K\in\mathcal{K}} \Big((Q, -\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}))_K - (\boldsymbol{\Sigma}, \boldsymbol{\sigma}-\boldsymbol{\Sigma})_K \Big) \tag{3.13}$$

$$= \sum_{K\in\mathcal{K}} \Big((\nabla Q, \boldsymbol{\sigma}-\boldsymbol{\Sigma})_K - (Q, \boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}))_{\partial K} - (\boldsymbol{\Sigma}, \boldsymbol{\sigma}-\boldsymbol{\Sigma})_K \Big) \tag{3.14}$$

$$= (\nabla Q - \boldsymbol{\Sigma}, \boldsymbol{\sigma}-\boldsymbol{\Sigma}) - \sum_{K\in\mathcal{K}} (Q, \boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}))_{\partial K} \tag{3.15}$$

$$\leq \|\nabla Q - \boldsymbol{\Sigma}\|_0^2 + \frac{1}{4}\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2 + \left| \sum_{K\in\mathcal{K}} (Q, \boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}))_{\partial K} \right|. \tag{3.16}$$

Using that $\boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})$ is continuous over element faces we can subtract an arbitrary function $v \in H^1(\Omega)$ in the term $\sum_{K\in\mathcal{K}}(Q, \boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}))_{\partial K} = \sum_{K\in\mathcal{K}}(Q-v, \boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}))_{\partial K}$. We then have the estimate

$$II \leq \|\nabla Q - \boldsymbol{\Sigma}\|_0^2 + \frac{1}{4}\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2 + \left| \inf_{v\in H^1(\Omega)} \sum_{K\in\mathcal{K}} (Q-v, \boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}))_{\partial K} \right|. \tag{3.17}$$

We now use the Cauchy-Schwartz inequality followed by the trace inequality,

$$\|\boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})\|_{-1/2,\partial K} \leq C(\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_{0,K} + h_K\|\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})\|_{0,K}) \tag{3.18}$$

$$= C(\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_{0,K} + h_K\|f+\nabla\cdot\boldsymbol{\Sigma}\|_{0,K}), \tag{3.19}$$

see [11], to estimate the sum in equation (3.17) as follows

$$\left| \inf_{v\in H^1(\Omega)} \sum_{K\in\mathcal{K}} (Q-v, \boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}))_{\partial K} \right| \tag{3.20}$$

$$\leq \inf_{v\in H^1(\Omega)} \sum_{K\in\mathcal{K}} \|Q-v\|_{1/2,\partial K}\|\boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})\|_{-1/2,\partial K} \tag{3.21}$$

$$\leq \inf_{v\in H^1(\Omega)} \left(\sum_{K\in\mathcal{K}} \|Q-v\|_{1/2,\partial K}^2 \right)^{1/2} \left(\sum_{K\in\mathcal{K}} \|\boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})\|_{-1/2,\partial K}^2 \right)^{1/2} \tag{3.22}$$

$$\leq C \inf_{v\in H^1(\Omega)} \left(\sum_{K\in\mathcal{K}} \|Q-v\|_{1/2,\partial K}^2 \right)^{1/2} \left(\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2 + \|h(f+\nabla\cdot\boldsymbol{\Sigma})\|_0^2 \right)^{1/2} \tag{3.23}$$

$$\leq C^2 \inf_{v\in H^1(\Omega)} \sum_{K\in\mathcal{K}} \|Q-v\|_{1/2,\partial K}^2 + \frac{1}{4}\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2 + \frac{1}{4}\|h(f+\nabla\cdot\boldsymbol{\Sigma})\|_0^2. \tag{3.24}$$

Together equation (3.17) and equation (3.20-3.24) gives a bound of the second term, II, in equation (3.7),

$$II \leq \|\nabla Q - \boldsymbol{\Sigma}\|_0^2 + \frac{1}{2}\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 + \frac{1}{4}\|h(f + \nabla \cdot \boldsymbol{\Sigma})\|_0^2 \\ + C^2 \inf_{v \in H^1(\Omega)} \sum_{K \in \mathcal{K}} \|Q - v\|_{1/2,\partial K}^2. \tag{3.25}$$

We combine equation (3.8) and equation (3.25) to get,

$$I + II \leq C\|\nabla Q - \boldsymbol{\Sigma}\|_0^2 + \frac{3}{4}\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 + C\|h(f + \nabla \cdot \boldsymbol{\Sigma})\|_0^2 \\ + C \inf_{v \in H^1(\Omega)} \sum_{K \in \mathcal{K}} \|Q - v\|_{1/2,\partial K}^2. \tag{3.26}$$

To estimate the last term on the right hand side in equation (3.26) we employ the technique of Lemma 4 in [3]. For completeness we include the details of the proof. We let $\mathcal{N}$ be the set of nodes in the mesh, $\{\phi\}_{i \in \mathcal{N}}$ piecewise linear base functions, $\omega_i = \text{supp}(\phi_i)$, CP_i continuous piecewise polynomials on ω_i , and $CP = \oplus_{i \in \mathcal{N}} \phi_i CP_i \in H^1(\Omega)$.

Using that $CP \subset H^1(\Omega)$ followed by the inverse inequality $\|Q - v\|_{1/2,\partial K}^2 \leq Ch_K^{-1}\|Q - v\|_{0,\partial K}^2$, which holds since both v and Q are piecewise polynomials, we get

$$\inf_{v \in H^1(\Omega)} \sum_{K \in \mathcal{K}} \|Q - v\|_{1/2,\partial K}^2 \leq \inf_{v \in CP} \sum_{K \in \mathcal{K}} \|Q - v\|_{1/2,\partial K}^2 \tag{3.27}$$

$$\leq C \inf_{v \in CP} \sum_{K \in \mathcal{K}} h_K^{-1}\|Q - v\|_{0,\partial K}^2. \tag{3.28}$$

We write $v = \sum_{i \in \mathcal{N}} \phi_i v_i \in CP$ and proceed with the estimate as follows

$$\|Q - v\|_{0,\partial K}^2 = \sum_{i \in \mathcal{N}} (Q - v, \phi_i(v_i - Q))_{\partial K} \tag{3.29}$$

$$\leq \sum_{i \in \mathcal{N}} \|\phi_i^{1/2}(Q - v)\|_{0,\partial K} \|\phi_i^{1/2}(v_i - Q)\|_{0,\partial K} \tag{3.30}$$

$$\leq \|Q - v\|_{0,\partial K} \left(\sum_{i \in \mathcal{N}} \|\phi_i^{1/2}(v_i - Q)\|_{0,\partial K}^2 \right)^{1/2}, \tag{3.31}$$

where we used that $\{\phi_i\}_{i \in \mathcal{N}}$ is a partition of unity. We have

$$\inf_{v \in H^1(\Omega)} \sum_{K \in \mathcal{K}} \|Q - v\|_{1/2,\partial K}^2 \leq \inf_{v \in CP} \sum_{i \in \mathcal{N}} \sum_{K \in \mathcal{K}} h_K^{-1}\|\phi_i^{1/2}(v_i - Q)\|_{0,\partial K}^2. \tag{3.32}$$

Further the following bound holds,

$$\inf_{v_i \in CP_i} \sum_{K \in \mathcal{K}} h_K^{-1} \|\phi_i^{1/2}(v_i - Q)\|_{0,\partial K}^2 \leq C \sum_{K \in \mathcal{K}} h_K^{-1} \|\phi_i^{1/2} [Q] \|_{0,\partial K}^2, \tag{3.33}$$

since the right hand side of equation (3.33) is zero when Q in continuous on ω_i so we may choose $v_i = Q|_{\omega_i}$ which means that the left hand side is also zero. The estimate follows from finite dimensionality and scaling. We end up with,

$$\inf_{v \in H^1(\Omega)} \sum_{K \in \mathcal{K}} \|Q - v\|_{1/2,\partial K}^2 \leq C \sum_{K \in \mathcal{K}} \sum_{i \in \mathcal{N}} h_K^{-1} \|\phi_i^{1/2} [Q] \|_{0,\partial K}^2 \tag{3.34}$$

$$= C \sum_{K \in \mathcal{K}} h_K^{-1} \| [Q] \|_{0,\partial K}^2, \tag{3.35}$$

again we use that $\{\phi_i\}_{i \in \mathcal{N}}$ is a partition of unity.

Combining equation (3.25) and (3.34) we get,

$$I + II \leq C\|\nabla Q - \boldsymbol{\Sigma}\|_0^2 + \frac{3}{4}\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 + C\|h(f + \nabla \cdot \boldsymbol{\Sigma})\|_0^2 + C \sum_{K \in \mathcal{K}} h_K^{-1} \| [Q] \|_{0,\partial K}^2. \tag{3.36}$$

Since $I + II = \|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2$ from equation (3.2-3.7) we just need to subtract $3/4\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2$ from both sides of equation (3.36) to prove the theorem. □

3.2 Estimate Based on Postprocessing

We now turn to the question of how to choose Q in Theorem 3.1. We know that choosing $Q = U$ results in a suboptimal estimate of the energy norm error, [4]. A natural idea is to choose Q to be a postprocessed version of U. There have been several works [8, 5, 17, 13] following Arnold and Brezzi [2], published in the mid eighties, on postprocessing methods where information from the calculated flux $\boldsymbol{\Sigma}$ is used to compute an improved approximation of u.

We focus on the method considered in Lovadina and Stenberg [13] and show that Theorem 3.1 directly gives the estimate presented in [13]. We denote the postprocessed version of U by U^*. To define U^* we introduce some notations. For all $K \in \mathcal{K}$ we let $P_{h,K} : L^2(\Omega) \to W_{h,K}$ be the L^2 projection onto $W_{h,K}$, where $W_{h,K}$ is the restriction of W_h onto K. Furthermore, we let $W_{h,K}^*$ denote the following spaces: $W_{h,K}^* = P_k(K)$ for RTN elements, $W_{h,K}^* = P_{k+1}(K)$ for BDM elements, and $W_{h,K}^* = P_{k+2}(K)$ for TH elements.

Definition 3.1 *(Postprocessing method) Find U^* such that $U^*|_K = U_K^* \in W_{h,K}^*$ where U_K^* is defined by*

$$P_{h,K} U_K^* = U_K, \tag{3.37}$$

and

$$(\nabla U^*, \nabla v)_K = (\boldsymbol{\Sigma}, \nabla v)_K \quad \textit{for all } v \in (I - P_{h,K}) W_{h,K}^*. \tag{3.38}$$

Proposition 3.1 *Given $f \in L^2(\Omega)$ it holds,*

$$\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 \leq C \sum_{K \in \mathcal{K}} \left(h_K^2 \|f + \nabla \cdot \boldsymbol{\Sigma}\|_{0,K}^2 + \|\boldsymbol{\Sigma} - \nabla U^*\|_{0,K}^2 + h_K^{-1} \| [U^*] \|_{0,\partial K}^2 \right), \tag{3.39}$$

where U^* is taken from Definition 3.1.

Proof. The proof follows directly from Theorem 3.1 with $Q = U^*$. □

Remark 3.1 In Corollary 2.6 in [13] the following a priori estimate of the error is presented for BDM and RTN elements,

$$\begin{aligned} \|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\| + \|U - U^*\|_1 &\leq C h^{k+1} |u|_{k+2} \quad \text{for BDM}, \\ \|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\| + \|U - U^*\|_1 &\leq C h^{k} |u|_{k+1} \quad \text{for BDM}, \end{aligned} \tag{3.40}$$

where $|\cdot|_k$ is the $H^k(\Omega)$ semi norm, see [1]. These estimates shows that the postprocessed function U^* gives optimal order estimates and indicates that the error estimators we present are of optimal order.

3.3 Estimate for Stabilized Methods

Here we extend our estimate to stabilized mixed methods, in particular, we consider the recent method presented in Masud and Hughes [14]. Stabilized methods are based on a modified weak formulation which yields a stable method for standard continuous piecewise polynomial approximations, e.g. piecewise linear functions for both displacement and flux.

The stabilized method of Masud and Hughes reads: find $\boldsymbol{\Sigma} \in \boldsymbol{V}_h$ and $U \in W_h$ such that,

$$(-\nabla \cdot \boldsymbol{\Sigma}, w) + (\boldsymbol{\Sigma}, \boldsymbol{v}) + (U, \nabla \cdot \boldsymbol{v}) - \frac{1}{2}(\boldsymbol{\Sigma} - \nabla U, \boldsymbol{v} + \nabla w) = (f, w), \tag{3.41}$$

for all $\boldsymbol{v} \in \boldsymbol{V}_h$ and $w \in W_h$. Applying the same ideas as in Theorem 3.1 to this stabilized method we obtain the following a posteriori error estimate. The argument may be modified to cover other stabilized methods such as the Galerkin least squares method.

Proposition 3.2 *For the approximate solution of equation (3.41) using continuous piecewise polynomials it holds,*

$$\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2 \leq C\sum_{K\in\mathcal{K}}\left(h_K^2\|f+\nabla\cdot\boldsymbol{\Sigma}\|_{0,K}^2+\|\boldsymbol{\Sigma}-\nabla U\|_{0,K}^2\right). \tag{3.42}$$

Proof. Using the same arguments as in equations (3.2-3.7) in the proof of Theorem 3.1, we obtain the following error representation formula,

$$\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2=(u-Q,f+\nabla\cdot\boldsymbol{\Sigma})+(Q,-\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}))-(\boldsymbol{\Sigma},\boldsymbol{\sigma}-\boldsymbol{\Sigma}). \tag{3.43}$$

Next, setting $\boldsymbol{v}=0$ in (3.41) we have the Galerkin orthogonality property

$$(f+\nabla\cdot\boldsymbol{\Sigma},w)=-\frac{1}{2}(\boldsymbol{\Sigma}-\nabla U,\nabla w), \tag{3.44}$$

for all $w\in W_h$. Subtracting the Scott-Zhang interpolant [6], $\pi_h(u-Q)$, of $u-Q$, using (3.44) followed by an interpolation estimate we get

$$\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2=(u-Q-\pi_h(u-Q),f+\nabla\cdot\boldsymbol{\Sigma})-\frac{1}{2}(\nabla\pi_h(u-Q),\boldsymbol{\Sigma}-\nabla U) \tag{3.45}$$

$$+(Q,-\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}))-(\boldsymbol{\Sigma},\boldsymbol{\sigma}-\boldsymbol{\Sigma}) \tag{3.46}$$

$$\leq C\sum_{K\in\mathcal{K}}\left(h_K^2\|f+\nabla\cdot\boldsymbol{\Sigma}\|_{0,K}^2+h_K^{-1}\|\,[Q]\,\|_{0,\partial K}^2\right)+C\|\boldsymbol{\Sigma}-\nabla Q\|_0^2 \tag{3.47}$$

$$+\frac{1}{2}\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2+\|\boldsymbol{\Sigma}-\nabla U\|_0\|\nabla\pi_h(u-Q)\|_0. \tag{3.48}$$

To get this estimate we also use arguments that are identical with the ones in the proof of Theorem 3.1. We choose $Q=U$. Since U is continuous the jump terms will vanish. We also use the stability of the interpolant π_h in $H^1(\Omega)$,

$$\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2\leq C\sum_{K\in\mathcal{K}}h_K^2\|f+\nabla\cdot\boldsymbol{\Sigma}\|_{0,K}^2+C\|\boldsymbol{\Sigma}-\nabla U\|_0^2 \tag{3.49}$$

$$+\frac{1}{2}\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2+C\|\boldsymbol{\Sigma}-\nabla U\|_0\|\boldsymbol{\sigma}-\nabla U\|_0 \tag{3.50}$$

$$\leq C\sum_{K\in\mathcal{K}}h_K^2\|f+\nabla\cdot\boldsymbol{\Sigma}\|_{0,K}^2+(C^2+C)\|\boldsymbol{\Sigma}-\nabla U\|_0^2 \tag{3.51}$$

$$+\frac{1}{2}\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2+\frac{1}{4}\|\boldsymbol{\sigma}-\nabla U\|_0^2. \tag{3.52}$$

But since $\|\boldsymbol{\sigma} - \nabla U\|_0 \leq \|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0 + \|\boldsymbol{\Sigma} - \nabla U\|_0$ we have,

$$\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 \leq C \sum_{K \in \mathcal{K}} h_K^2 \|f + \nabla \cdot \boldsymbol{\Sigma}\|_{0,K}^2 + C\|\boldsymbol{\Sigma} - \nabla U\|_0^2 + \frac{3}{4}\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2, \quad (3.53)$$

so the proposition follows immediately after subtracting $3/4\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2$ from both sides. □

References

[1] R. A. Adams, *Sobolev spaces,* volume 65 of Pure and Applied Mathematics, Academic Press, New York, 1975.

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

[3] R. Becker, P. Hansbo, and M. G. Larson, *Energy norm a posteriori error estimation for discontinuous Galerkin methods,* Comput. Methods Appl. Mech. Engrg., 192, 723-733, 2003.

[4] D. Braess and R. Verfürt, *A posteriori error estimators for the Raviart-Thomas element,* SIAM J. Numer. Anal., 33, 2431-2444, 1996.

[5] J. H. Bramble and J. Xu, *A local post-processing technique for improving the accuracy in mixed finite-element approximations,* SIAM J. Numer. Anal., 26, 1267-1275, 1989.

[6] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods,* Springer Verlag, 1994.

[7] F. Brezzi, J. Douglas Jr., R. Durán, and M. Fortin, *Mixed finite elements for second order elliptic problems in three variables,* Numer. Math., 51, 237-250, 1987.

[8] F. Brezzi, J. Douglas Jr., and D. L. Marini, *Two families of mixed finite elements for second order elliptic problems,* Numer. Math., 47, 217-235, 1985.

[9] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods,* Springer Verlag, 1991.

[10] C. Carstensen, *A posteriori error estimate for the mixed finite element method,* Math. Comp., 66, 465-476, 1997.

[11] V. Girault, P.-A. Raviart, *Finite element approximation of the Navier-Stokes equation,* Springer Verlag, Berlin, 1979.

[12] P. Hood and C. Taylor, *A numerical solution of the Navier-Stokes equations using the finite element techniques,* Comp. and Fluids, 1, 73-100, 1973.

[13] C. Lovadina and R. Stenberg, *Energy norm a posteriori error estimates for mixed finite element methods,* Helsinki University of Technology. Research Report A473. September 2004.

[14] A. Masud and T. J. R. Hughes, *A stabilized finite element method for Darcy flow,* Comput. Methods Appl. Mech. Engrg., 191, 4341-4370, 2002.

[15] J.-C. Nédélec, *A new family of mixed finite elements in* $\mathbf{R}^3$*,* Numer. Math., 50, 57-81, 1986.

[16] P. Raviart and J. Thomas, *A mixed finite element method for second order elliptic problems,* in Mathematical Aspects of the Finite Element Method. Lecture Notes in Math. 606, Springer-Verlag, 292-315, 1977.

[17] R. Stenberg, *Postprocessing schemes for some mixed finite elements,* RAIRO Modél. Math. Anal. Numér., 25, 151-167, 1991.

A Mixed Adaptive Variational Multiscale Method with Applications in Oil Reservoir Simulation

Axel Målqvist*

August 19, 2005

Abstract

We present a mixed adaptive variational multiscale method for solving elliptic second order problems. This work is an extension of the adaptive variational multiscale method (AVMS), introduced by Larson and Målqvist [15, 16, 17], to a mixed formulation. The method is based on a particular splitting into coarse and fine scales together with a systematic technique for approximation of the fine scale part based on solution of decoupled localized subgrid problems. We present the mixed AVMS method and derive a posteriori error estimates for both linear functionals and the energy norm. Based on the estimates we propose an adaptive algorithm for automatic tuning of critical discretization parameters. Finally, we present numerical examples on a two dimensional slice of an oil reservoir.

1 Introduction

We consider the Poisson equation on mixed form with positive diffusion coefficient $a \in L^\infty(\Omega)$:

$$\begin{cases} \frac{1}{a}\boldsymbol{\sigma} - \nabla u = 0 & \text{in } \Omega, \\ -\nabla \cdot \boldsymbol{\sigma} = f & \text{in } \Omega, \\ n \cdot \boldsymbol{\sigma} = 0 & \text{on } \Gamma. \end{cases} \tag{1.1}$$

We assume the integral over the polygonal domain Ω of the right hand side to be zero, $\int_\Omega f\,dx = 0$, in order to get a well posed problem with a solution $u \in H^1(\Omega)/\mathbf{R}$ and $\boldsymbol{\sigma} \in H(\mathrm{div};\Omega)$, see [3] for definitions of these spaces. The

*Graduate Research Assistant, Department of Mathematics, Chalmers University of Technology, Göteborg, axel@math.chalmers.se.

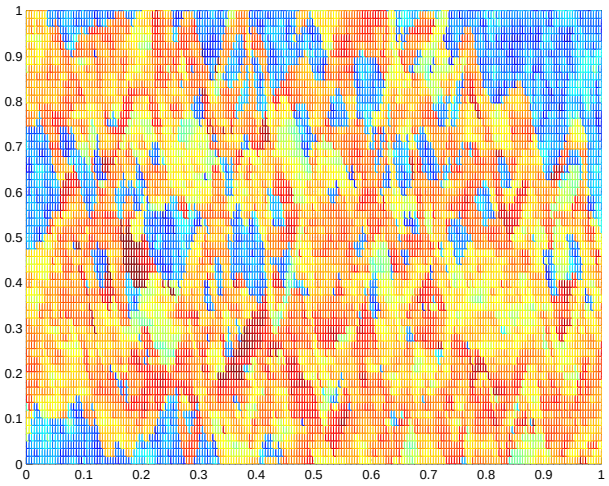


Figure 1: Two dimensional slice of the permeability a in an oil reservoir (in log scale) taken from the tenth SPE comparative solution project $\max a / \min a = 8.8 \cdot 10^6$.

boundary of Ω is denoted Γ. Our main focus is to develop a method for solving this problem in the case when a has fine scale features that can not be resolved by a single mesh, see Figure 1 for a typical coefficient a with microstructure.

Previous Work. Various multiscale methods have been developed in recent years. Hou and Wu [12] presented a method in 1997 based on homogenization theory. This method has been improved and extended by Aarnes and Lie et. al. [1, 2] and successfully applied to oil reservoir problems.

The variational multiscale method was introduced in 1995 by Hughes [13, 14]. Here the spaces are divided into a coarse and a fine part and then analytical methods are used to approximate the effect of the fine scales on the coarse scale. See also the work by Arbogast [4] for mixed formulations.

In [15, 16, 17] we develop a new multiscale method for the standard formulation of Poisson's equation with diffusion coefficient a:

$$\begin{cases} -\nabla \cdot a\nabla u = f & \text{in } \Omega, \\ u = 0 & \text{on } \Gamma. \end{cases} \tag{1.2}$$

We also derive a posteriori error estimates both in the energy norm and for linear functionals and based on these estimates we propose and implement adaptive algorithms. In [19] we prove optimal energy norm a posteriori error estimates for standard mixed finite element approximations of (1.1), with

$a = 1$, such as the Raviart-Thomas [21, 20], BDM [8], and Taylor-Hood [11] elements. Together [15, 16, 17, 19] provides the theoretic starting point to develop a mixed adaptive variational multiscale method for the Poisson equation.

New Contributions. In this paper we present the mixed adaptive variational method. We also derive a posteriori error estimates for both linear functionals and the energy norm of the error and based on the estimate we formulate an adaptive algorithm that automatically tunes the parameters in the method according to the error estimate. This is a very important feature of the method. We apply the adaptive method on an oil reservoir problem and obtain very promising results.

Outline. The remainder of the paper is organized as follows: in Section 2 we present the method; we discuss implementation issues in Section 3; in Section 4 we derive the a posteriori error estimators used in Section 5 to formulate an adaptive algorithm; and finally in Section 6 we present numerical results.

2 The Mixed Variational Multiscale Method

We start by deriving the variational formulation of equation (1.1) by multiplying the first equation with a test function $\boldsymbol{v} \in \boldsymbol{V} = \{\boldsymbol{v} \in H(\mathrm{div}; \Omega) : n \cdot \boldsymbol{v} = 0 \text{ on } \Gamma\}$, integrate over the domain Ω, and integrate by parts. We also multiply the second equation by a test function $w \in W = L^2(\Omega)$ and integrate over Ω. The weak form reads: find $\boldsymbol{\sigma} \in \boldsymbol{V}$ and $u \in W$ such that,

$$\begin{cases} (\frac{1}{a}\boldsymbol{\sigma}, \boldsymbol{v}) + (u, \nabla \cdot \boldsymbol{v}) = 0, \\ \quad -(\nabla \cdot \boldsymbol{\sigma}, w) = (f, w), \end{cases} \tag{2.1}$$

for all $\boldsymbol{v} \in \boldsymbol{V}$ and $w \in W$, where $(v, w) = (v, w)_\Omega = \int_\Omega v\, w\, dx$. Since we focus on problems that features fine scale behavior it is natural to split the spaces $\boldsymbol{V} = \boldsymbol{V}_c \oplus \boldsymbol{V}_f$ and $W = W_c \oplus W_f$ into a discrete coarse part that we can resolve using a standard finite element method on a single mesh and a fine part that needs to be taken care of in a non-standard way.

By this argument we end up with the variational multiscale formulation, see [13] for an overview, of equation (1.1): find $\boldsymbol{\sigma}_c \in \boldsymbol{V}_c$, $\boldsymbol{\sigma}_f \in \boldsymbol{V}_f$, $u_c \in W_c$, and $u_f \in W_f$ such that,

$$\begin{cases} (\frac{1}{a}\boldsymbol{\sigma}_c, \boldsymbol{v}_c) + (\frac{1}{a}\boldsymbol{\sigma}_f, \boldsymbol{v}_c) + (u_c, \nabla \cdot \boldsymbol{v}_c) + (u_f, \nabla \cdot \boldsymbol{v}_c) = 0, \\ \quad -(\nabla \cdot \boldsymbol{\sigma}_c, w_c) - (\nabla \cdot \boldsymbol{\sigma}_f, w_c) = (f, w_c), \\ (\frac{1}{a}\boldsymbol{\sigma}_f, \boldsymbol{v}_f) + (u_f, \nabla \cdot \boldsymbol{v}_f) = -(\frac{1}{a}\boldsymbol{\sigma}_c, \boldsymbol{v}_f) - (u_c, \nabla \cdot \boldsymbol{v}_f), \\ \quad -(\nabla \cdot \boldsymbol{\sigma}_f, w_f) = (f, w_f) + (\nabla \cdot \boldsymbol{\sigma}_c, w_f), \end{cases} \tag{2.2}$$

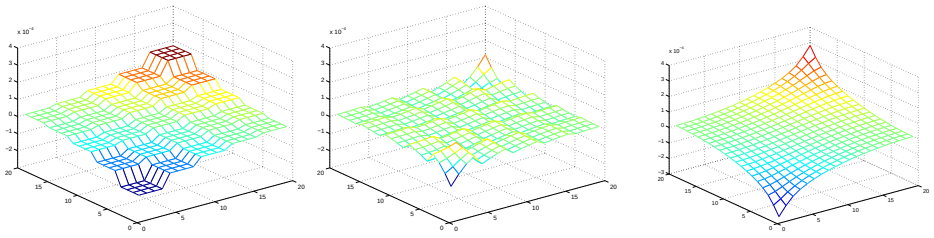


Figure 2: The pressure u_c, u_f, and $u = u_c + u_f$.

for all $\boldsymbol{v}_c \in \boldsymbol{V}_c$, $\boldsymbol{v}_f \in \boldsymbol{V}_f$, $w_c \in W_c$, and $w_f \in W_f$. We note that the fine scale equations are driven by the coarse scale residual multiplied by the fine scale test functions. In Figure 2 we illustrate u_c, u_f, and u in an example where $a = 1$ and f is -1 in the lower left corner and 1 in the upper right corner.

2.1 Splitting into Coarse and Fine Scales

There are many possible ways of splitting the spaces $\boldsymbol{V}$ and W into coarse and fine parts. This choice is fundamental since the splitting determines in what sense the approximate coarse scale part of given by the method will minimize the error. It is obvious that this aspect is of great importance when deriving a method.

For example if we would like to have a method that aims at nodal exactness on the coarse mesh nodes we would choose a nodal based hierarchical basis. This works since V_f in this case is zero in the coarse nodes and we know that we have not made any approximations hence V_c will be nodal exact in coarse nodes, see also [15, 16, 17]. However since we are interested in problems with strong fine scale behavior it is more natural to approximate some kind of mean value of the solution locally.

In this paper we will use the lowest order Raviart-Thomas elements on rectangles together with piecewise constants for the coarse spaces. For the fine scale we use the natural hierarchical basis. This choice makes sense in terms of the optimized output quantity which will, for the pressure, be a mean value (L^2 projection) over coarse elements rather then a point value and for the fluxes we get the meanvalues over the faces.

We also get another nice feature which simplified both the analysis and the numerics. Some terms in equations (2.2) will actually vanish using this split,

$$(w_c, \nabla \cdot \boldsymbol{v}_f) = \sum_{K \in \mathcal{K}} (w_c, \nabla \cdot \boldsymbol{v}_f)_K = \sum_{K \in \mathcal{K}} w_c^K \int_{\partial K} \boldsymbol{n} \cdot \boldsymbol{v}_f \, dx = 0, \qquad (2.3)$$

where w_c^K is the constant value w_c has at coarse element K and $\mathcal{K} = \cup K$ is the set of coarse elements on Ω. We also have,

$$(w_f, \nabla \cdot \boldsymbol{v}_c) = \sum_K (w_f, \nabla \cdot \boldsymbol{v}_c)_K = \sum_K \nabla \cdot \boldsymbol{v}_c^K \int_K w_f \, dx = 0, \tag{2.4}$$

where $\nabla \cdot \boldsymbol{v}_c^K$ is the constant vector $\nabla \cdot \boldsymbol{v}_c$ at coarse element K. Equation (2.3) and (2.4) holds for all $\boldsymbol{v}_c \in \boldsymbol{V}_c$, $\boldsymbol{v}_f \in \boldsymbol{V}_f$, $w_c \in W_c$, and $w_f \in W_f$.

If we for the moment only modify the coarse scale equations (2.2) transforms to: find $\boldsymbol{\sigma}_c \in \boldsymbol{V}_c$, $\boldsymbol{\sigma}_f \in \boldsymbol{V}_f$, $u_c \in W_c$, and $u_f \in W_f$ such that,

$$\left\{ \begin{array}{r} (\frac{1}{a}\boldsymbol{\sigma}_c, \boldsymbol{v}_c) + (\frac{1}{a}\boldsymbol{\sigma}_f, \boldsymbol{v}_c) + (u_c, \nabla \cdot \boldsymbol{v}_c) = 0, \\ -(\nabla \cdot \boldsymbol{\sigma}_c, w_c) = (f, w_c), \\ (\frac{1}{a}\boldsymbol{\sigma}_f, \boldsymbol{v}_f) + (u_f, \nabla \cdot \boldsymbol{v}_f) = -(\frac{1}{a}\boldsymbol{\sigma}_c, \boldsymbol{v}_f) - (u_c, \nabla \cdot \boldsymbol{v}_f), \\ -(\nabla \cdot \boldsymbol{\sigma}_f, w_f) = (f, w_f) + (\nabla \cdot \boldsymbol{\sigma}_c, w_f), \end{array} \right. \tag{2.5}$$

for all $\boldsymbol{v}_c \in \boldsymbol{V}_c$, $\boldsymbol{v}_f \in \boldsymbol{V}_f$, $w_c \in W_c$, and $w_f \in W_f$.

2.2 Decoupling Fine Scale Equations.

Since the fine scale equations are to computationally expensive to solve in one global computation we seek a way to decouple them. We construct two partitions of unity related to the base functions in analogy with the earlier papers [15, 16, 17]. We let ψ_i be piecewise constant on the coarse mesh such that $\sum_{i \in \mathcal{N}} \psi_i = 1$, see Figure 3 (left). Further we let $\boldsymbol{\phi}_i$ be the lowest order Raviart-Thomas basis functions on the coarse mesh, see Figure 3 (right). For the Raviart-Thomas space in two spatial dimensions we form a matrix partition of unity $\{\phi_i\}_{i \in \mathcal{N}}$ in the following way,

$$\boldsymbol{\phi}_i = \begin{bmatrix} \phi_i^x \\ \phi_i^y \end{bmatrix}, \qquad \phi_i = \begin{bmatrix} \phi_i^x & 0 \\ 0 & \phi_i^y \end{bmatrix}, \tag{2.6}$$

and in the same way in three dimensions. We note that $\sum_{i \in \mathcal{N}} \phi_i = I$, where I is the identity matrix. We include these partitions of unity next to the fine scale test functions and end up with a decoupled problem. We introduce $\boldsymbol{\sigma}_{f,i} \in \boldsymbol{V}_f$ and $u_{f,i} \in W_f$ such that $\boldsymbol{\sigma}_c$, u_c, $\boldsymbol{\sigma}_f = \sum_{i \in \mathcal{N}} \boldsymbol{\sigma}_{f,i}$, and $u_f = \sum_{i \in \mathcal{N}} u_{f,i}$ solves:

$$\left\{ \begin{array}{r} (\frac{1}{a}\boldsymbol{\sigma}_c, \boldsymbol{v}_c) + (\frac{1}{a}\boldsymbol{\sigma}_f, \boldsymbol{v}_c) + (u_c, \nabla \cdot \boldsymbol{v}_c) = 0, \\ -(\nabla \cdot \boldsymbol{\sigma}_c, w_c) = (f, w_c), \\ (\frac{1}{a}\boldsymbol{\sigma}_{f,i}, \boldsymbol{v}_f) + (u_{f,i}, \nabla \cdot \boldsymbol{v}_f) = -(\frac{1}{a}\boldsymbol{\sigma}_c, \phi_i \boldsymbol{v}_f) - (u_c, \nabla \cdot (\phi_i \boldsymbol{v}_f)), \\ -(\nabla \cdot \boldsymbol{\sigma}_{f,i}, w_f) = (f, \psi_i w_f) + (\nabla \cdot \boldsymbol{\sigma}_c, \psi_i w_f), \end{array} \right. \tag{2.7}$$

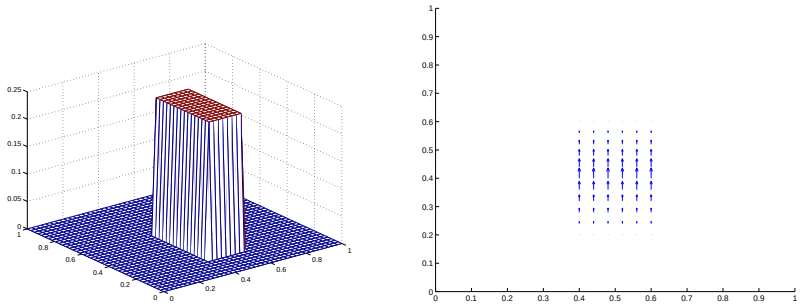


Figure 3: Elements of the two partitions of unity, to the left pressure ψ_i, and to the right flux $\boldsymbol{\phi}_i$.

for all $\boldsymbol{v}_c \in \boldsymbol{V}_c$, $\boldsymbol{v}_f \in \boldsymbol{V}_f$, $w_c \in W_c$, and $w_f \in W_f$. Again we can drop two terms since,

$$(w_c, \nabla \cdot (\phi_i \boldsymbol{v}_f)) = \sum_{K \in \mathcal{K}} w_c^K \int_{\partial K} \boldsymbol{n} \cdot (\phi_i \boldsymbol{v}_f)\, dx = 0, \tag{2.8}$$

and,

$$(\nabla \cdot \boldsymbol{v}_c, \psi_i w_f) = \sum_K \nabla \cdot \boldsymbol{v}_c^K \psi_K \int_K w_f\, dx = 0. \tag{2.9}$$

Thus we get the following simplified problem: find $\boldsymbol{\sigma}_c \in \boldsymbol{V}_c$, $u_c \in W_c$, $\boldsymbol{\sigma}_{f,i} \in \boldsymbol{V}_f$, and $u_{f,i} \in W_f$ such that,

$$\begin{cases} (\frac{1}{a}\boldsymbol{\sigma}_c, \boldsymbol{v}_c) + (\frac{1}{a}\boldsymbol{\sigma}_f, \boldsymbol{v}_c) + (u_c, \nabla \cdot \boldsymbol{v}_c) = 0, \\ -(\nabla \cdot \boldsymbol{\sigma}_c, w_c) = (f, w_c), \\ (\frac{1}{a}\boldsymbol{\sigma}_{f,i}, \boldsymbol{v}_f) + (u_{f,i}, \nabla \cdot \boldsymbol{v}_f) = -(\frac{1}{a}\boldsymbol{\sigma}_c, \phi_i \boldsymbol{v}_f), \\ -(\nabla \cdot \boldsymbol{\sigma}_{f,i}, w_f) = (f, \psi_i w_f), \end{cases} \tag{2.10}$$

for all $\boldsymbol{v}_c \in \boldsymbol{V}_c$, $\boldsymbol{v}_f \in \boldsymbol{V}_f$, $w_c \in W_c$, and $w_f \in W_f$.

To get an idea of how typical fine scale solutions $\boldsymbol{\sigma}_{f,i}$ and $u_{f,i}$ can look like we give an example in Figure 4 where $a = 1$ and $f = -1$ in the lower left corner and $f = 1$ in the upper right corner. The basis functions in the right hand side are chosen in the middle of the domain.

2.3 Solving Local Neumann Problems on Patches

We let $\boldsymbol{V}_c = \boldsymbol{V}_H$ be the lowest order Raviart-Thomas finite element space and $W_c = W_H$ be the space of piecewise constants on a mesh $\cup K = \mathcal{K}$ with mesh function H defined as a piecewise constant function equal to the diameter of the current element. Since $\boldsymbol{V}_f$ and W_f are infinite dimensional

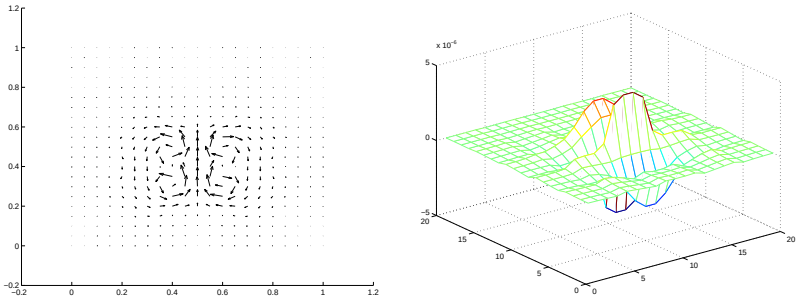


Figure 4: The local solutions $\boldsymbol{\sigma}_{f,i}$ and $u_{f,i}$

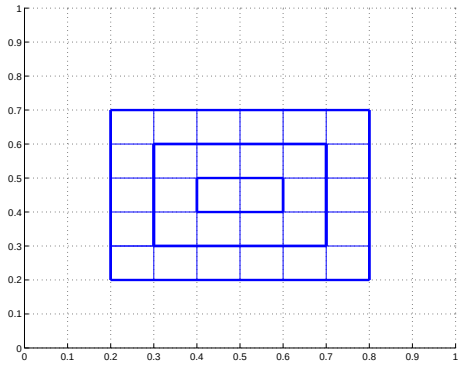


Figure 5: One, two, and three layer patches made out of coarse elements.

they also need to be discretized. We introduce a new mesh function $h \ll H$ individually for the decoupled problems by refining the coarse mesh once or several times. These meshes resolves the fine scale behavior but are to rich to solve a discrete system of equations on. To get around this problem we introduce patches on which we solve decoupled fine scale equations instead of solving them on the entire domain.

Since the spaces $\boldsymbol{V}_H$ and W_H are discrete we seek approximate solutions with coarse part in these spaces $\boldsymbol{\Sigma}_c \in \boldsymbol{V}_H$ and $U_c \in W_H$. For the approximation of the fine scale solutions $\boldsymbol{\Sigma}_{f,i}$ and $U_{f,i}$ we introduce corresponding solution spaces, $\boldsymbol{V}_h(\overline{\omega_i})$ and $W_h(\overline{\omega_i})$ where $\{\omega_i\}$ are domains (patches) such that,

$$\text{supp}(\boldsymbol{\phi}_i) = \text{supp}(\psi_i) \subset \omega_i \subset \Omega, \tag{2.11}$$

see Figure 5 for three different sizes of patches, all including the support of the basis functions (the two coarse elements in the center of the picture in Figure (5)). The fine scale spaces $\boldsymbol{V}_h(\overline{\omega_i})$ and $W_h(\overline{\omega_i})$ are constructed on the

patch ω_i as a hierarchical extension of the spaces $\boldsymbol{V}_H$ and W_H restricted to the patch. Since they are hierarchical spaces all $\boldsymbol{v} \in \boldsymbol{V}_h(\overline{\omega_i})$ fulfil $\int_F \boldsymbol{n}\cdot\boldsymbol{v}\, ds = 0$ on each coarse face F and all $w \in W_h(\overline{\omega_i})$ fulfil $\int_K w\, dx = 0$ on each coarse element K. The bars over the domains in the fine scale spaces symbolizes that we have not yet forced any boundary conditions on the spaces. Since we aim at solving local Neumann problems on patches we want to include the boundary condition in the definition of the spaces. The final fine scale spaces we will use are the following,

$$\boldsymbol{V}_h(\omega_i) = \{\boldsymbol{v} \in \boldsymbol{V}_h(\overline{\omega_i}) \ : \ \boldsymbol{n}\cdot\boldsymbol{v} = 0 \text{ on } \partial\omega_i\}, \tag{2.12}$$

and

$$W_h(\omega_i) = W_h(\overline{\omega_i}). \tag{2.13}$$

The piecewise constant space is not affected by the boundary conditions as seen above. We sometimes refer to $\boldsymbol{V}_h(\omega_i)$ and $W_h(\omega_i)$ as slice spaces since they are just the fine scale part of a classic finite element spaces.

We will frequently refer to layers in the rest of this paper. Layers are a measure of how large the patches are. One layer will be the two coarse element on which the coarse Raviart-Thomas base function has its support, two layers are these two coarse element together with all coarse element that neighbors the first two and so on. If we use rectangular elements the n-layer patches will also be rectangular and the higher n is the larger the patch becomes. See Figure 5 for one, two, and, three layer patches.

We are now ready to present the method. The finite element formulation reads: find $\boldsymbol{\Sigma}_c \in \boldsymbol{V}_H, \boldsymbol{\Sigma}_{f,i} \in \boldsymbol{V}_h(\omega_i)$, $U_c \in W_H$, and $U_{f,i} \in W_h(\omega_i)$ such that

$$\begin{cases} (\frac{1}{a}\boldsymbol{\Sigma}_c, \boldsymbol{v}_c) + (\frac{1}{a}\boldsymbol{\Sigma}_f, \boldsymbol{v}_c) + (U_c, \nabla\cdot\boldsymbol{v}_c) = 0, \\ -(\nabla\cdot\boldsymbol{\Sigma}_c, w_c) = (f, w_c), \\ (\frac{1}{a}\boldsymbol{\Sigma}_{f,i}, \boldsymbol{v}_f) + (U_{f,i}, \nabla\cdot\boldsymbol{v}_f) = -(\frac{1}{a}\boldsymbol{\Sigma}_c\phi_i, \boldsymbol{v}_f), \\ -(\nabla\cdot\boldsymbol{\Sigma}_{f,i}, w_f) = (f, w_f\psi_i), \end{cases} \tag{2.14}$$

for all $\boldsymbol{v}_c \in \boldsymbol{V}_H$, $\boldsymbol{v}_f \in \boldsymbol{V}_h(\omega_i)$, $w_c \in W_H$, $w_f \in W_h(\omega_i)$, and $i \in \mathcal{N}$.

For computational reasons we are also going to modify the decoupling presented in equation (2.10) slightly in order to get a simplified version of the method. We will replace $\boldsymbol{\sigma}_c\phi_i$ by $\sigma_c^i\boldsymbol{\phi}_i$, where $\boldsymbol{\sigma}_c = \sum_{i\in\mathcal{N}}\sigma_c^i\boldsymbol{\phi}_i$, and $\sigma_c^i \in \mathbf{R}$. On discrete form the simplified method reads: find $\boldsymbol{\Sigma}_c = \sum_{i\in\mathcal{N}}\Sigma_c^i\boldsymbol{\phi}_i \in \boldsymbol{V}_H$, $U_c \in W_H$, $\boldsymbol{\Sigma}_{f,i} \in \boldsymbol{V}_h(\omega_i)$, and $U_{f,i} \in W_h(\omega_i)$ such that,

$$\begin{cases} (\frac{1}{a}\boldsymbol{\Sigma}_c, \boldsymbol{v}_c) + (\frac{1}{a}\boldsymbol{\Sigma}_f, \boldsymbol{v}_c) + (U_c, \nabla\cdot\boldsymbol{v}_c) = 0, \\ -(\nabla\cdot\boldsymbol{\sigma}_c, w_c) = (f, w_c), \\ (\frac{1}{a}\boldsymbol{\Sigma}_{f,i}, \boldsymbol{v}_f) + (U_{f,i}, \nabla\cdot\boldsymbol{v}_f) = -(\frac{1}{a}\Sigma_c^i\boldsymbol{\phi}_i, \boldsymbol{v}_f), \\ -(\nabla\cdot\boldsymbol{\Sigma}_{f,i}, w_f) = (f, \psi_i w_f), \end{cases} \tag{2.15}$$

for all $\boldsymbol{v}_c \in \boldsymbol{V}_H$, $\boldsymbol{v}_f \in \boldsymbol{V}_h(\omega_i)$, $w_c \in W_c$, and $w_f \in W_h(\omega_i)$. Note that we only use a slightly different partition of unity. The error estimates of the two methods (2.14) and (2.15) will be the same with the only difference that we replace $\boldsymbol{\Sigma}_c\phi_i$ by $\Sigma_c^i\boldsymbol{\phi}_i$ in all equations. We present the error analysis for the original method (2.14).

In the next section we discuss implementation issues and here we use the simplified version of (2.15) the method in order to make the presentation more accessible to the reader.

3 Implementation

3.1 Iterative Method

There are primarily two ways to implement the method. The first one that comes to mind is perhaps an iterative approach. We simply assume $\boldsymbol{\Sigma}_f = 0$ in equation (2.15) and solve the coarse problem. This gives the lowest order Raviart-Thomas approximate solution on the coarse grid. Then we use this solution in the right hand side of the fine scale equations and calculate a new version of $\boldsymbol{\Sigma}_f = \sum_i \boldsymbol{\Sigma}_{f,i}$. We include the new approximation of $\boldsymbol{\Sigma}_f$ into the coarse scale equations and solve them again and so on. This method yields a sequence of solutions $\{\boldsymbol{\Sigma}_c^j\}_{j=1}^m$ and $\{\boldsymbol{\Sigma}_f^j\}_{j=1}^m$ which hopefully converges to the desired solution $\boldsymbol{\Sigma}$. However in this paper we will not analyze or use this approach. We will instead use a direct method.

3.2 Direct Method

The aim here is to directly get the converged solution by modifying the matrices and vectors of the linear system of equations. When solving the Poisson equation on mixed form using Raviart-Thomas elements we end up with the following system of equations,

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B}^T & \mathbf{O} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{b} \end{bmatrix}, \tag{3.1}$$

where

$$\begin{aligned} \mathbf{A}_{ij} &= (\tfrac{1}{a}\boldsymbol{\phi}_j, \boldsymbol{\phi}_i), \\ \mathbf{B}_{ij} &= (\psi_j, \nabla \cdot \boldsymbol{\phi}_i), \\ \mathbf{b}_j &= (f, \psi_j). \end{aligned} \tag{3.2}$$

To get a similar system of equations using the multiscale method we need to express $\boldsymbol{\Sigma}_f$ in terms of $\boldsymbol{\Sigma}_c$. To do this we introduce the following two fine scale help problems. The first one reads: find $\boldsymbol{\xi}_i \in \boldsymbol{V}_h(\omega_i)$ and $\zeta_i \in W_h(\omega_i)$ such that

$$\begin{cases} (\tfrac{1}{a}\boldsymbol{\xi}_i, \boldsymbol{v}_f) + (\zeta_i, \nabla \cdot \boldsymbol{v}_f) = -(\tfrac{1}{a}\boldsymbol{\phi}_i, \boldsymbol{v}_f), \\ -(\nabla \cdot \boldsymbol{\xi}_i, w_f) = 0, \end{cases} \tag{3.3}$$

for all $\boldsymbol{v}_f \in \boldsymbol{V}_h(\omega_i)$, and $w_f \in W_h(\omega_i)$ and $i \in \mathcal{N}$. Here f is replaced by 0 and Σ_c^i is replaced by 1 compared to the fine part of equation (2.15). The second problem reads: find $\boldsymbol{\beta}_i \in \boldsymbol{V}_h(\omega_i)$ and $\rho_i \in W_h(\omega_i)$ such that

$$\begin{cases} (\frac{1}{a}\boldsymbol{\beta}_i, \boldsymbol{v}_f) + (\rho_i, \nabla \cdot \boldsymbol{v}_f) = 0, \\ -(\nabla \cdot \boldsymbol{\beta}_i, w_f) = (f\psi_i, w_f), \end{cases} \tag{3.4}$$

for all $\boldsymbol{v}_f \in \boldsymbol{V}_h(\omega_i)$, and $w_f \in W_h(\omega_i)$ and $i \in \mathcal{N}$. Here instead Σ_c^i is replaced by 0. Since all equations are linear we conclude that,

$$\boldsymbol{\Sigma}_{f,i} = \Sigma_c^i \boldsymbol{\xi}_i + \boldsymbol{\beta}_i, \tag{3.5}$$

so if we let $\boldsymbol{\beta} = \sum_i \boldsymbol{\beta}_i$,

$$\boldsymbol{\Sigma}_f = \sum_{i \in \mathcal{N}} \Sigma_c^i \boldsymbol{\xi}_i + \boldsymbol{\beta}. \tag{3.6}$$

We are now ready to go back to the coarse part of equation (2.15) and replace $\boldsymbol{\Sigma}_f$ by the expression derived in equation (3.6),

$$\begin{cases} (\frac{1}{a}\sum_j \Sigma_c^j \boldsymbol{\phi}_j, \boldsymbol{\phi}_i) + (\frac{1}{a}(\sum_j \Sigma_c^j \boldsymbol{\xi}_j + \boldsymbol{\beta}), \boldsymbol{\phi}_i) + (\sum_k U_c^k \psi_k, \nabla \cdot \boldsymbol{\phi}_i) = 0, \\ -(\nabla \cdot \sum_j \Sigma_c^j \boldsymbol{\phi}_j, \psi_i) = (f, \psi_k). \end{cases} \tag{3.7}$$

If we write this system on matrix form in analogy with equation (3.1) we get,

$$\begin{bmatrix} \mathbf{A} + \mathbf{T} & \mathbf{B} \\ -\mathbf{B}^T & \mathbf{O} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{d} \\ \mathbf{b} \end{bmatrix}, \tag{3.8}$$

where

$$\begin{aligned} \mathbf{T}_{ij} &= (\tfrac{1}{a}\boldsymbol{\xi}_j, \boldsymbol{\phi}_i), \\ \mathbf{d}_j &= -(\tfrac{1}{a}\boldsymbol{\beta}, \boldsymbol{\phi}_j). \end{aligned} \tag{3.9}$$

3.3 Is the Method Locally Conservative?

If we assume that the local meshes coincide for all patches we can proceed with the following calculation,

$$(-\nabla \cdot \boldsymbol{\Sigma}, w_f) = (-\nabla \cdot \boldsymbol{\Sigma}_f, w_f) \tag{3.10}$$

$$= \sum_{i \in \mathcal{N}} (-\nabla \cdot \boldsymbol{\Sigma}_{f,i}, w_f) \tag{3.11}$$

$$= \sum_{i \in \mathcal{N}} (f\psi_i, w_f) \tag{3.12}$$

$$= (f, w_f), \tag{3.13}$$

where we use equation (2.15). This shows that the method is conservative using this assumption. If the meshes differ we still have a conservative

method on the refinement level where the meshes coincide. If we for example have one mesh refinement in all patches but one where we have two mesh refinements compared to the coarse mesh we have a conservative method for one refinement over the whole domain.

3.4 Is the Method Parallel?

The fine scale problems described in equation (3.3) and (3.4) are totally decoupled and can be solved without any communication between each other. After the solutions $\boldsymbol{\xi}_i$ and $\boldsymbol{\beta}_i$ have been calculated we can compute the entries in the T matrix and d vector directly in the local calculation. We just multiply all test functions that are non zero on the current patch with $\boldsymbol{\xi}_i$ and $\boldsymbol{\beta}_i$. This means that we never need to actually construct and save any global data on the fine reference space. This is an important feature of the method. In many cases it is impossible to form the total fine scale solution for storage reasons. The method is parallel, the fine scale patches only communicates with the coarse scale and only in one direction.

3.5 A Symmetric Method

Since the new stiffness matrix $A + T$ is built up of elements $(A + T)_{ij} = (\frac{1}{a}\boldsymbol{\phi}_j, \boldsymbol{\phi}_i + \boldsymbol{\xi}_i)$ we see directly that it is not symmetric anymore. This is a bit unfortunate and may cause instabilities in the method. However, if we for a moment assume the patches to be equal to the entire domain $\omega_i = \Omega$ and that we use the same resolution on all local problems, which means $\boldsymbol{\Sigma}$ is the standard Raviart-Thomas solution on the fine mesh we note, using the third part of equation (2.15) summed over all i, that,

$$(\frac{1}{a}\boldsymbol{\Sigma}, \boldsymbol{\xi}_i) = -(U_f, \nabla \cdot \boldsymbol{\xi}_i) = 0, \tag{3.14}$$

since $(\nabla \cdot \boldsymbol{\xi}_i, w_f) = 0$ for all $w_f \in \boldsymbol{V}_h(\omega_i) = \boldsymbol{V}_h(\Omega)$. So in the limit it is equivalent to let $(A + T)_{ij} = (\frac{1}{a}\boldsymbol{\phi}_j + \boldsymbol{\xi}_j, \boldsymbol{\phi}_i + \boldsymbol{\xi}_i)$. We also note that independent of the patches we always have,

$$(\psi_j, \nabla \cdot \boldsymbol{\xi}_i) = 0, \tag{3.15}$$

since $\boldsymbol{\xi}_i$ is a fine scale function. This lead us to formulate a modified version of the method:

$$\begin{bmatrix} (\mathbf{A} + \mathbf{T})^{sym} & \mathbf{B} \\ -\mathbf{B}^T & \mathbf{O} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{d}^{sym} \\ \mathbf{b} \end{bmatrix}, \tag{3.16}$$

where

$$\begin{aligned}
&(\mathbf{A}+\mathbf{T})^{sym}_{ij} = (\tfrac{1}{a}\boldsymbol{\phi}_j + \boldsymbol{\xi}_j, \boldsymbol{\phi}_i + \boldsymbol{\xi}_i),\\
&\mathbf{B}_{ij} = (\psi_j, \nabla\cdot(\boldsymbol{\phi}_i + \boldsymbol{\xi}_i)) = (\psi_j, \nabla\cdot\boldsymbol{\phi}_i),\\
&\mathbf{b} = (f, \psi_j),\\
&\mathbf{d}^{sym}_j = -(\tfrac{1}{a}\boldsymbol{\phi}_j + \boldsymbol{\xi}_j, \boldsymbol{\beta}).
\end{aligned} \tag{3.17}$$

Comparing equation (3.16) and (3.17) with equation (3.1) and (3.2) we note that they differ in $\mathbf{d}$ and that the old base functions for the flux are replaced by new ones,

$$\boldsymbol{\phi}_i := \boldsymbol{\phi}_i + \boldsymbol{\xi}_i. \tag{3.18}$$

This makes the stiffness matrix $(\mathbf{A}+\mathbf{T})^{sym}$ symmetric and positive definite. The symmetric multiscale method and the standard multiscale method described in this paper converge to the same result, when using meshes that coincide, as the patch size increases, however they will give different results for small patches. If the standard method seems to need very big patches to work the symmetric method may be an alternative.

The size of the patches will play a crucial role for both methods. The only thing we know so far is that we will get the correct solution on the fine mesh if we let all $\omega_i = \Omega$. This result is not interesting from a practical point of view. So let us first discuss why we expect a good approximation from this method at all, when using a few layer patches. The right hand side of the fine scale equations has the same support as the elements in the partition of unity. We also solve the fine scale equations on the slice spaces $\boldsymbol{V}_h(\omega_i)$ which makes the solution decay much faster since $\int_F \boldsymbol{n}\cdot\boldsymbol{\Sigma}_{f,i}\,dx = 0$ for all coarse faces F. This forces the solution to be very localized. This leads us to believe that we can expect a good approximation even if we decrease the size of the patch around the current coarse edge on which we solve the fine scale problem.

To be able to give a more solid statement of the accuracy of the method we need good a posteriori error element-indicators. These will be used to formulate an adaptive strategy that will increase the accuracy. We want to make a method that automatically tunes the patch size and fine scale mesh size in order to give us an optimal solution given the amount of time we are ready to spend.

In the next section we derive such error indicators for linear functionals and in energy norm, further on we present adaptive algorithms for automatic improvement of the solution. Adaptivity is a fundamental part of this method.

4 A Posteriori Error Estimate

4.1 Postprocessing of $U_{f,i}$

In this section we introduce functions $U^*_{f,i}$ that are post processed (improved) versions of $U_{f,i}$. It is known from [18] and [19] that in order to get optimal a posteriori error estimates U has to be replaced by a postprocessed version U^* when computing the residuals for mixed methods using a richer approximation space for $\mathbf{\Sigma}$ than U. For example, this is the case for the lowest order Raviart-Thomas together with piecewise constant elements. There are different methods for computing U^*, see [18] and the references therein. The basic idea is to use the displacement U and the flux $\mathbf{\Sigma}$ to determine U^* in a richer space.

In the multiscale setting only the fine part of U need to be modified. In the error analysis in the next section we will not assume any particular choice of $U^*_{f,i}$ other then $U^*_{f,i} \in \bigoplus_{K\in\mathcal{K}} P_1(K)$, where $P_1(K)$ is bilinear functions on element K. However, in practise we compute $U^*_{f,i} \in P_1(K)$ using the following equations

$$(U^*_{f,i}, v) = (U_{f,i}, v) \quad \text{for all } v \in P_0(K), \tag{4.1}$$

$$(a\nabla U^*_{f,i}, \nabla w) = (\mathbf{\Sigma}_{f,i} + \mathbf{\Sigma}_c\phi_i, \nabla w) \quad \text{for all } w \in (I - P_K)P_1(K), \tag{4.2}$$

where P_K is the local $L^2(K)$ onto constants.

Finally, we also need a technical definition to state the error representation formula. We let F_i be the coarse face on which $|\boldsymbol{\phi}_i| = 1$ i.e.

$$F_i = \{x : |\boldsymbol{\phi}_i| = 1\}. \tag{4.3}$$

In Figure 3 F_i is the line between $(x, y) = (0.4, 0.4)$ and $(x, y) = (0.6, 0.4)$. We next define the auxiliary function $U_{c,i}$.

Definition 4.1 *For each fine element $K \in \mathcal{K}(\omega_i)$ we let U_c^K be the constant interior value of U_c on K and define $U_{c,i}$ on ∂K as,*

$$\begin{cases} U_{c,i} = U_c^K \quad \textit{when } x \in F_i, \\ U_{c,i} = 0 \quad \textit{otherwise.} \end{cases} \tag{4.4}$$

Thus $U_{c,i}$ is defined on all faces of every fine scale element K. It takes the value zero on all faces except F_i. On F_i it takes either value of the discontinuous function U_c depending on which side of F_i element K is on. For example in Figure 3, $U_{c,i}$ would have one constant value below F_i and an other constant value above F_i.

4.2 Error Representation for Linear Functionals

We introduce the following dual problem,

$$\begin{cases} \frac{1}{a}\boldsymbol{\chi} - \nabla\eta = \boldsymbol{\omega}, \\ -\nabla\cdot\boldsymbol{\chi} = 0. \end{cases} \tag{4.5}$$

Using the dual problem we can derive an error representation formula for a linear functional.

Proposition 4.1 *For arbitrary $U^*_{f,i} \in \bigoplus_{K\in\mathcal{K}} P_1(K)$ it holds,*

$$\begin{aligned}(\boldsymbol{\sigma}-\boldsymbol{\Sigma},\boldsymbol{\omega}) = \sum_{i\in\mathcal{N}}(-\frac{1}{a}(\boldsymbol{\Sigma}_c\phi_i+\boldsymbol{\Sigma}_{f,i})+\nabla U^*_{f,i},\boldsymbol{\chi}) \\ -\sum_{i\in\mathcal{N}}\sum_{K\in\mathcal{K}(\omega_i)}(U^*_{f,i}+U_{c,i},\boldsymbol{n}\cdot\boldsymbol{\chi})_{\partial K\setminus\partial\omega_i} \\ -\sum_{i\in\mathcal{N}}(U^*_{f,i},\boldsymbol{n}\cdot\boldsymbol{\chi})_{\partial\omega_i\setminus\Gamma} \\ -\sum_{i\in\mathcal{N}}(f\psi_i+\nabla\cdot(\boldsymbol{\Sigma}_c\phi_i+\boldsymbol{\Sigma}_{f,i}),\eta-\pi_c\eta-\pi_{f,i}\eta).\end{aligned} \tag{4.6}$$

Proof. We use the dual problem to get,

$$(\boldsymbol{\sigma}-\boldsymbol{\Sigma},\boldsymbol{\omega}) = (\boldsymbol{\sigma}-\boldsymbol{\Sigma},\frac{1}{a}\boldsymbol{\chi}-\nabla\eta) \tag{4.7}$$

$$= (\frac{1}{a}\boldsymbol{\sigma},\boldsymbol{\chi}) - (\frac{1}{a}\boldsymbol{\Sigma},\boldsymbol{\chi}) - (\boldsymbol{\sigma}-\boldsymbol{\Sigma},\nabla\eta) \tag{4.8}$$

$$= -(u,\nabla\cdot\boldsymbol{\chi}) - (\frac{1}{a}\boldsymbol{\Sigma},\boldsymbol{\chi}) + (\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}),\eta) \tag{4.9}$$

$$= -(\frac{1}{a}\boldsymbol{\Sigma},\boldsymbol{\chi}) + (\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}),\eta), \tag{4.10}$$

where we use the boundary conditions to eliminate the boundary terms, $\boldsymbol{n}\cdot\boldsymbol{\sigma} = \boldsymbol{n}\cdot\boldsymbol{\Sigma} = 0$. We note from the second equation in (2.14) together with equation (2.3) that,

$$(f+\nabla\cdot\boldsymbol{\Sigma},\pi_c\eta) = (f+\nabla\cdot\boldsymbol{\Sigma}_c,\pi_c\eta) = 0. \tag{4.11}$$

Next, we continue the main calculation using (4.11),

$$(\boldsymbol{\sigma}-\boldsymbol{\Sigma},\boldsymbol{\omega}) = -(\frac{1}{a}\boldsymbol{\Sigma},\boldsymbol{\chi}) + (\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}),\eta-\pi_c\eta) \tag{4.12}$$

$$\begin{aligned}= \sum_{i\in\mathcal{N}} -(\frac{1}{a}(\boldsymbol{\Sigma}_c\phi_i+\boldsymbol{\Sigma}_{f,i}),\boldsymbol{\chi}) \\ -\sum_{i\in\mathcal{N}}(f\psi_i+\nabla\cdot(\boldsymbol{\Sigma}_c\phi_i+\boldsymbol{\Sigma}_{f,i}),\eta-\pi_c\eta).\end{aligned} \tag{4.13}$$

In analogy with [19] we are now ready to add the term $-(U, \nabla \cdot \boldsymbol{\chi}) = 0$ to equation (4.13). We let $\mathcal{K}_H(\omega_i)$ be the set of coarse elements on patch ω_i. We have,

$$0 = -(U, \nabla \cdot \boldsymbol{\chi}) \tag{4.14}$$

$$= -\sum_{i \in \mathcal{N}} \left((U^*_{f,i}, \nabla \cdot \boldsymbol{\chi}) + (U_c, \nabla \cdot (\phi_i \boldsymbol{\chi})) \right) \tag{4.15}$$

$$= \sum_{i \in \mathcal{N}} \left((\nabla U^*_{f,i}, \boldsymbol{\chi}) - \sum_{K \in \mathcal{K}(\omega_i)} (U^*_{f,i}, \boldsymbol{n} \cdot \boldsymbol{\chi})_{\partial K} - \sum_{K \in \mathcal{K}_H(\omega_i)} (U_c, \boldsymbol{n} \cdot (\phi_i \boldsymbol{\chi}))_{\partial K} \right) \tag{4.16}$$

using integration by parts and the fact that ∇U_c is zero in the interior of coarse elements. We would like to have the two last terms of equation (4.16) on the same form. To accomplish this we need to turn to the last term.

Since we use rectangular elements $(U_c, \boldsymbol{n} \cdot (\phi_i \boldsymbol{\chi}))_{\partial K}$ will be zero on all but one face in the coarse mesh. This is due to orthogonality between $\boldsymbol{n}$ and $\phi_i \boldsymbol{\chi}$ and that ϕ_i is only non zero on two coarse elements. The one face where $(U_c, \boldsymbol{n} \cdot (\phi_i \boldsymbol{\chi}))_{\partial K} \neq 0$ is F_i, see equation (4.3). Since $(U_c, \boldsymbol{n} \cdot (\phi_i \boldsymbol{\chi}))_{\partial K} = (U_c, \boldsymbol{n} \cdot \boldsymbol{\chi})_{\partial K}$ on F_i we use the auxiliary function $U_{c,i}$ defined earlier to conclude that, $(U_c, \boldsymbol{n} \cdot (\phi_i \boldsymbol{\chi}))_{\partial K} = (U_{c,i}, \boldsymbol{n} \cdot \boldsymbol{\chi})_{\partial K}$.

Continuing from equation (4.14) we get,

$$0 = \sum_{i \in \mathcal{N}} \left((\nabla U^*_{f,i}, \boldsymbol{\chi}) - \sum_{K \in \mathcal{K}(\omega_i)} (U^*_{f,i}, \boldsymbol{n} \cdot \boldsymbol{\chi})_{\partial K} - \sum_{K \in \mathcal{K}_H(\omega_i)} (U_c, \boldsymbol{n} \cdot (\phi_i \boldsymbol{\chi}))_{\partial K} \right) \tag{4.17}$$

$$= \sum_{i \in \mathcal{N}} \left((\nabla U^*_{f,i}, \boldsymbol{\chi}) - \sum_{K \in \mathcal{K}(\omega_i)} (U^*_{f,i}, \boldsymbol{n} \cdot \boldsymbol{\chi})_{\partial K} - \sum_{K \in \mathcal{K}(\omega_i)} (U_{c,i}, \boldsymbol{n} \cdot \boldsymbol{\chi})_{\partial K} \right) \tag{4.18}$$

$$= \sum_{i \in \mathcal{N}} \left((\nabla U^*_{f,i}, \boldsymbol{\chi}) - \sum_{K \in \mathcal{K}(\omega_i)} (U^*_{f,i} + U_{c,i}, \boldsymbol{n} \cdot \boldsymbol{\chi})_{\partial K} \right). \tag{4.19}$$

Adding this term to equation (4.13) gives,

$$\begin{aligned} (\boldsymbol{\sigma} - \boldsymbol{\Sigma}, \boldsymbol{\omega}) = \sum_{i \in \mathcal{N}} & (-\frac{1}{a}(\boldsymbol{\Sigma}_c \phi_i + \boldsymbol{\Sigma}_{f,i}) + \nabla U^*_{f,i}, \boldsymbol{\chi})_{\omega_i} \\ & - \sum_{i \in \mathcal{N}} \sum_{K \in \mathcal{K}(\omega_i)} (U^*_{f,i} + U_{c,i}, \boldsymbol{n} \cdot \boldsymbol{\chi})_{\partial K} \\ & - \sum_{i \in \mathcal{N}} (f \psi_i + \nabla \cdot (\boldsymbol{\Sigma}_c \phi_i + \boldsymbol{\Sigma}_{f,i}), \eta - \pi_c \eta - \pi_{f,i} \eta). \end{aligned} \tag{4.20}$$

Finally, splitting the second term in equation (4.20) into an interior and a boundary part proves the proposition. □

The error representation formula (4.6) can be used to create an adaptive algorithm as it stands. The dual problem needs to be solved numerically. The more effort we put into the calculation of the solution to the dual problem the better control of the error we get. We may also proceed on the calculation using estimates to get abound of the error.

4.3 Error Estimates

In this paper we are primarily interested in deriving a posteriori error estimates which are easy to implement and scale properly with the error so that they can serve as a basis for adaptive algorithms. We are thus not interested in tracking constants C independent of the mesh size.

Theorem 4.1 *For arbitrary $U^*_{f,i} \in \bigoplus_{K\in\mathcal{K}} P_1(K)$ it holds,*

$$
\begin{aligned}
|(\boldsymbol{\sigma}-\boldsymbol{\Sigma},\boldsymbol{\omega})| \leq & \sum_{i\in\mathcal{N}} C\|\sqrt{a}\|_{L^\infty(\omega_i)}\|\nabla U^*_{f,i} - \frac{1}{a}(\boldsymbol{\Sigma}_c\phi_i + \boldsymbol{\Sigma}_{f,i})\|_{\omega_i}\|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\|_{\omega_i} \qquad (4.21)\\
& + \sum_{i\in\mathcal{N}} C\|\sqrt{a}\|_{L^\infty(\omega_i)}\left(\sum_{K\in\mathcal{K}(\omega_i)} h^{-1}\|[U^*_{f,i}+U_{c,i}]\|^2_{\partial K\setminus\partial\omega_i}\right)^{1/2}\|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\|_{\omega_i}\\
& + \sum_{i\in\mathcal{N}} C\|\sqrt{a}\|_{L^\infty(\omega_i)}\|h^{-1/2}U^*_{f,i}\|_{\partial\omega_i\setminus\Gamma}\|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\|_{\omega_i}\\
& + \sum_{i\in\mathcal{N}} C\|\sqrt{a}\|_{\tilde{L}^\infty(\omega_i)}\|\frac{h}{a}(f\psi_i + \nabla\cdot(\boldsymbol{\Sigma}_c\phi_i+\boldsymbol{\Sigma}_{f,i}))\|_{\omega_i}\|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\|_{\omega_i}\\
& + \sum_{i\in\mathcal{N}} C\|\sqrt{a}\|_{\tilde{L}^\infty(\omega_i)}\|\frac{h}{a}(f\psi_i + \nabla\cdot(\boldsymbol{\Sigma}_c\phi_i+\boldsymbol{\Sigma}_{f,i}))\|_{\omega_i}\|\sqrt{a}\boldsymbol{\omega}\|_{\omega_i},
\end{aligned}
$$

where we use the notation,

$$
\|\sqrt{a}\|_{\tilde{L}^\infty(\omega_i)} = \max_{K\in\mathcal{K}(\omega_i)} \|a\|_{L^\infty(K)}\|\frac{1}{\sqrt{a}}\|_{L^\infty(K)}. \qquad (4.22)
$$

Proof. Start from equation (4.20) we have

$$(\boldsymbol{\sigma}-\boldsymbol{\Sigma},\boldsymbol{\omega}) = \sum_{i\in\mathcal{N}}(-\frac{1}{a}(\boldsymbol{\Sigma}_c\phi_i+\boldsymbol{\Sigma}_{f,i})+\nabla U^*_{f,i},\boldsymbol{\chi}) \tag{4.23}$$
$$-\sum_{i\in\mathcal{N}}\sum_{K\in\mathcal{K}(\omega_i)}(U^*_{f,i}+U_{c,i},\boldsymbol{n}\cdot\boldsymbol{\chi})_{\partial K}$$
$$-\sum_{i\in\mathcal{N}}(f\psi_i+\nabla\cdot(\boldsymbol{\Sigma}_c\phi_i+\boldsymbol{\Sigma}_{f,i}),\eta-\pi_c\eta-\pi_{f,i}\eta)$$
$$= \mathrm{I}+\mathrm{II}+\mathrm{III}. \tag{4.24}$$

We thus have three terms that we will treat separately. We use Cauchy-Schwartz and Hölder's inequalities to get a bound of the first term,

$$\mathrm{I} = \sum_{i\in\mathcal{N}}(-\frac{1}{a}(\boldsymbol{\Sigma}_c\phi_i+\boldsymbol{\Sigma}_{f,i})+\nabla U^*_{f,i},\boldsymbol{\chi}) \tag{4.25}$$
$$\leq \sum_{i\in\mathcal{N}}\|\nabla U^*_{f,i}-\frac{1}{a}(\boldsymbol{\Sigma}_c\phi_i+\boldsymbol{\Sigma}_{f,i})\|_{\omega_i}\|\boldsymbol{\chi}\|_{\omega_i} \tag{4.26}$$
$$\leq \sum_{i\in\mathcal{N}}\|\sqrt{a}\|_{L^\infty(\omega_i)}\|\nabla U^*_{f,i}-\frac{1}{a}(\boldsymbol{\Sigma}_c\phi_i+\boldsymbol{\Sigma}_{f,i})\|_{\omega_i}\|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\|_{\omega_i}. \tag{4.27}$$

Next turning to the second term II we note that we can subtract a continuous function equal to zero on the boundary of the patch from $U^*_{f,i}+U_{c,i}$ in the second term II of equation (4.23). For $v\in H^1_0(\omega_i)=\{v : v\in H^1(\omega_i)$ such that $v=0$ on $\partial\omega_i\}$ we have

$$\sum_{K\in\mathcal{K}(\omega_i)}-(U^*_{f,i}+U_{c,i},\boldsymbol{n}\cdot\boldsymbol{\chi})_{\partial K} = \inf_{v\in H^1_0(\omega_i)}\sum_{K\in\mathcal{K}(\omega_i)}-(U^*_{f,i}+U_{c,i}-v,\boldsymbol{n}\cdot\boldsymbol{\chi})_{\partial K} \tag{4.28}$$
$$\leq \sum_{K\in\mathcal{K}(\omega_i)}\|U^*_{f,i}+U_{c,i}-v\|_{1/2,\partial K\setminus\Gamma}\|\boldsymbol{n}\cdot\boldsymbol{\chi}\|_{-1/2,\partial K} \tag{4.29}$$
$$\leq \left(C\sum_{K\in\mathcal{K}(\omega_i)}h^{-1}\|[U^*_{f,i}+U_{c,i}]\|^2_{\partial K\setminus\Gamma}\right)^{1/2}\left(\sum_{K\in\mathcal{K}(\omega_i)}\|\boldsymbol{n}\cdot\boldsymbol{\chi}\|^2_{-1/2,\partial K}\right)^{1/2}. \tag{4.30}$$

Here we used the estimate,

$$\inf_{v\in H^1_0(\omega_i)}\sum_{K\in\mathcal{K}(\omega_i)}\|U^*_{f,i}+U_{c,i}-v\|^2_{1/2,\partial K}\leq C\sum_{K\in\mathcal{K}(\omega_i)}h^{-1}\|[U^*_{f,i}+U_{c,i}]\|^2_{\partial K}, \tag{4.31}$$

from Lemma 4 in [5] with $V_{g_D} = H_0^1(\omega_i)$ and $\boldsymbol{n} \cdot \boldsymbol{\chi} = 0$ on Γ which allows us to replace $U_{f,i}^* + U_{c,i} - v$ by zero on Γ. To treat the second term of equation (4.30) we use the following trace inequality together with the observation $\nabla \cdot \boldsymbol{\chi} = 0$,

$$\|\boldsymbol{n} \cdot \boldsymbol{\chi}\|_{-1/2,\partial K} \leq C(\|\boldsymbol{\chi}\|_K + h\|\nabla \cdot \boldsymbol{\chi}\|_K) = C\|\boldsymbol{\chi}\|_K, \tag{4.32}$$

see [10]. Continuing the calculation in (4.30) we get,

$$\sum_{K \in \mathcal{K}(\omega_i)} - (U_{f,i}^* + U_{c,i}, \nabla \cdot \boldsymbol{\chi}) \tag{4.33}$$

$$\leq C \left(\sum_{K \in \mathcal{K}(\omega_i)} h^{-1} \|[U_{f,i}^* + U_{c,i}]\|_{\partial K \setminus \Gamma}^2 \right)^{1/2} \left(\sum_{K \in \mathcal{K}(\omega_i)} \|\boldsymbol{\chi}\|_K^2 \right)^{1/2} \tag{4.34}$$

$$\leq C\|\sqrt{a}\|_{L^\infty(\omega_i)} \left(\sum_{K \in \mathcal{K}(\omega_i)} h^{-1} \|[U_{f,i}^* + U_{c,i}]\|_{\partial K \setminus \Gamma}^2 \right)^{1/2} \|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\|_{\omega_i}. \tag{4.35}$$

Summation over $i \in \mathcal{N}$ gives,

$$II \leq C \sum_{i \in \mathcal{N}} \|\sqrt{a}\|_{L^\infty(\omega_i)} \left(\sum_{K \in \mathcal{K}(\omega_i)} h^{-1} \|[U_{f,i}^* + U_{c,i}]\|_{\partial K \setminus \Gamma}^2 \right)^{1/2} \|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\|_{\omega_i}. \tag{4.36}$$

We can now split the interior and boundary contributions from equation (4.36) to get the terms in the Proposition.

Finally, we consider the last term, III. This term is standard and straight forward to estimate. We simply use the interpolation estimate,

$$\|\eta - \pi_c \eta - \pi_{f,i}\eta\|_K \leq Ch\|\nabla \eta\|_K, \tag{4.37}$$

and get,

$$\text{III} \leq \sum_{i \in \mathcal{N}} \sum_{K \in \mathcal{K}(\omega_i)} Ch\|(f\psi_i + \nabla \cdot (\boldsymbol{\Sigma}_c \phi_i + \boldsymbol{\Sigma}_{f,i}))\|_K \|\nabla \eta\|_K \tag{4.38}$$

$$\leq \sum_{i \in \mathcal{N}} C\|\sqrt{a}\|_{\tilde{L}^\infty(\omega_i)} \|\frac{h}{a}(f\psi_i + \nabla \cdot (\boldsymbol{\Sigma}_c \phi_i + \boldsymbol{\Sigma}_{f,i}))\|_{\omega_i} \|\sqrt{a}\nabla \eta\|_{\omega_i} \tag{4.39}$$

$$\leq \sum_{i \in \mathcal{N}} C\|\sqrt{a}\|_{\tilde{L}^\infty(\omega_i)} \|\frac{h}{a}(f\psi_i + \nabla \cdot (\boldsymbol{\Sigma}_c \phi_i + \boldsymbol{\Sigma}_{f,i}))\|_{\omega_i} \|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\|_{\omega_i} \tag{4.40}$$

$$+ \sum_{i \in \mathcal{N}} C\|\sqrt{a}\|_{\tilde{L}^\infty(\omega_i)} \|\frac{h}{a}(f\psi_i + \nabla \cdot (\boldsymbol{\Sigma}_c \phi_i + \boldsymbol{\Sigma}_{f,i}))\|_{\omega_i} \|\sqrt{a}\boldsymbol{\omega}\|_{\omega_i},$$

where

$$\|\sqrt{a}\|_{\tilde{L}^\infty(\omega_i)} = \max_{K\in\mathcal{K}(\omega_i)} \|a\|_{L^\infty(K)} \|\frac{1}{\sqrt{a}}\|_{L^\infty(K)}. \tag{4.41}$$

If we finally combine equation (4.23, 4.27, 4.36, and 4.40) the theorem follows immediately. □

This estimate is of the correct form for adaptivity. For each patch we have element indicators that can be decrease either by decreasing h or increasing the patch size. The first, second, and fourth term decays when h decreases. Since $U^*_{f,i}$ will decay rapidly on the boundary of the patch when the patch size increases we can get a rapid decay in the third term in the error estimate as well. Another observation from the estimate is that if, $\boldsymbol{\omega} = \nabla w$, i.e. $\boldsymbol{\omega}$ can be written as a gradient we will get trivial solutions to the dual problem, $\eta = w$ and $\boldsymbol{\chi} = 0$ which will simplify the estimate significantly. This is not very surprising since,

$$(\boldsymbol{\sigma} - \boldsymbol{\Sigma}, \boldsymbol{\omega}) = (\boldsymbol{\sigma} - \boldsymbol{\Sigma}, \nabla w) = -(\nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\Sigma}), w) = (f + \nabla \cdot \boldsymbol{\Sigma}, w). \tag{4.42}$$

Finally, we use Theorem 4.1 to get an error estimate in the energy norm,

$$\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_a = (\frac{1}{a}\boldsymbol{\sigma} - \boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma})^{1/2}. \tag{4.43}$$

Corollary 4.1 *For arbitrary $U^*_{f,i} \in \bigoplus_{K\in\mathcal{K}} P_1(K)$ it holds,*

$$\begin{aligned}
\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_a \leq & \left(\sum_{i\in\mathcal{N}} C\|\sqrt{a}\|^2_{L^\infty(\omega_i)} \|\nabla U^*_{f,i} - \frac{1}{a}(\boldsymbol{\Sigma}_c\phi_i + \boldsymbol{\Sigma}_{f,i})\|^2_{\omega_i} \right)^{1/2} \\
& + \left(\sum_{i\in\mathcal{N}} C\|\sqrt{a}\|^2_{L^\infty(\omega_i)} \sum_{K\in\mathcal{K}(\omega_i)} h^{-1} \|[U^*_{f,i} + U_{c,i}]\|^2_{\partial K\setminus\partial\omega_i} \right)^{1/2} \\
& + \left(\sum_{i\in\mathcal{N}} C\|\sqrt{a}\|^2_{L^\infty(\omega_i)} \|h^{-1/2} U^*_{f,i}\|^2_{\partial\omega_i\setminus\Gamma} \right)^{1/2} \\
& + \left(\sum_{i\in\mathcal{N}} C\|\sqrt{a}\|^2_{\tilde{L}^\infty(\omega_i)} \|\frac{h}{a}(f\psi_i + \nabla\cdot(\boldsymbol{\Sigma}_c\phi_i + \boldsymbol{\Sigma}_{f,i}))\|^2_{\omega_i} \right)^{1/2}.
\end{aligned} \tag{4.44}$$

Proof. We will simply use Cauchy-Schwartz repeatedly on the result of

Theorem 4.1,

$$|(\boldsymbol{\sigma}-\boldsymbol{\Sigma},\boldsymbol{\omega})| \leq \left(\sum_{i\in\mathcal{N}} C\|\sqrt{a}\|^2_{L^\infty(\omega_i)}\|\nabla U^*_{f,i} - \frac{1}{a}(\boldsymbol{\Sigma}_c\phi_i + \boldsymbol{\Sigma}_{f,i})\|^2_{\omega_i}\right)^{1/2} \|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\| \tag{4.45}$$

$$+ \left(\sum_{i\in\mathcal{N}} C\|\sqrt{a}\|^2_{L^\infty(\omega_i)} \sum_{K\in\mathcal{K}(\omega_i)} h^{-1}\|[U^*_{f,i}+U_{c,i}]\|^2_{\partial K\setminus\partial\omega_i}\right)^{1/2} \|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\|$$

$$+ \left(\sum_{i\in\mathcal{N}} C\|\sqrt{a}\|^2_{L^\infty(\omega_i)}\|h^{-1/2}U^*_{f,i}\|^2_{\partial\omega_i\setminus\Gamma}\right)^{1/2} \|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\|$$

$$+ \left(\sum_{i\in\mathcal{N}} C\|\sqrt{a}\|^2_{\tilde{L}^\infty(\omega_i)}\|\frac{h}{a}(f\psi_i + \nabla\cdot(\boldsymbol{\Sigma}_c\phi_i+\boldsymbol{\Sigma}_{f,i}))\|^2_{\omega_i}\right)^{1/2} \|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\|$$

$$+ \left(\sum_{i\in\mathcal{N}} C\|\sqrt{a}\|^2_{\tilde{L}^\infty(\omega_i)}\|\frac{h}{a}(f\psi_i + \nabla\cdot(\boldsymbol{\Sigma}_c\phi_i+\boldsymbol{\Sigma}_{f,i}))\|^2_{\omega_i}\right)^{1/2} \|\sqrt{a}\boldsymbol{\omega}\|.$$

Next we note that if we multiply equation (4.5) by $\boldsymbol{\chi}$ and integrate over the domain,

$$\|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\|^2 = (\boldsymbol{\omega},\boldsymbol{\chi}) + (\nabla\eta,\boldsymbol{\chi}) = (\sqrt{a}\boldsymbol{\omega},\frac{1}{\sqrt{a}}\boldsymbol{\chi}) - (\eta,\nabla\cdot\boldsymbol{\chi}) \leq \|\sqrt{a}\boldsymbol{\omega}\|\|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\|. \tag{4.46}$$

Hence $\|\frac{1}{\sqrt{a}}\boldsymbol{\chi}\| \leq \|\sqrt{a}\boldsymbol{\omega}\|$. We now simply choose,

$$\boldsymbol{\omega} = \frac{1}{a}(\boldsymbol{\sigma}-\boldsymbol{\Sigma}), \tag{4.47}$$

and the Corollary follows immediately. □

5 The Adaptive Algorithm

We will base our adaptive strategy on the energy norm estimate presented in Corollary 4.1. The same ideas may be used in the duality based case. The parameters we will tune to get an improved solution are the individual mesh size on each patch and the individual size of each patch. We do not discuss how the coarse mesh is chosen. The idea is that we have reached a point where we can not afford a richer coarse space and need to consider parallel methods. The coarse mesh will in practise probably be constructed

using a standard adaptive algorithm for a single mesh, see [9, 6] for such methods.

We will use the error indicators from Corollary 4.1 to refine and extend the patches. The refinements does not need to be uniform since we can calculate indicators for individual fine element on the patch. However in this section we will present a simple algorithm that uses different but uniform meshes on each patch.

We have the following four error indicators if we ignore the constants,

$$\begin{cases} \|\sqrt{a}\|^2_{L^\infty(\omega_i)} \|\nabla U^*_{f,i} - \frac{1}{a}(\mathbf{\Sigma}_c \phi_i + \mathbf{\Sigma}_{f,i})\|^2_{\omega_i}, \\ \|\sqrt{a}\|^2_{L^\infty(\omega_i)} \sum_{K \in \mathcal{K}(\omega_i)} h^{-1} \|[U^*_{f,i} + U_{c,i}]\|^2_{\partial K \setminus \partial \omega_i}, \\ \|\sqrt{a}\|^2_{L^\infty(\omega_i)} \|h^{-1/2} U^*_{f,i}\|^2_{\partial \omega_i \setminus \Gamma}, \\ \|\sqrt{a}\|^2_{\tilde{L}^\infty(\omega_i)} \|\frac{h}{a}(f\psi_i + \nabla \cdot (\mathbf{\Sigma}_c \phi_i + \mathbf{\Sigma}_{f,i}))\|^2_{\omega_i}. \end{cases} \tag{5.1}$$

We are mainly interested in creating an adaptive algorithm that automatically improves the solution in an iterative fashion based on an error estimate that scales correctly in the parameters of interest. The main goal is therefore not to calculate a good approximation of the error. This means that we are more interested in how the four indicators compare to each other than of their absolute value. But when we look at the four term we immediately see that the fourth term has a different a dependent term in front of it. However this is not a big issue since if the a coefficient is constant on the fine mesh, and we can assume that it is close to constant since we have resolved the coefficient with the fine mesh, we have the following identity,

$$\|\sqrt{a}\|^2_{L^\infty(\omega_i)} = \max_{K \in \mathcal{K}(\omega_i)} \|a\|^2_{L^\infty(K)} \|\frac{1}{\sqrt{a}}\|^2_{L^\infty(K)} := \|\sqrt{a}\|^2_{\tilde{L}^\infty(\omega_i)}. \tag{5.2}$$

With this result in mind we are ready to present four indicators that will be used in the adaptive algorithm,

$$\begin{cases} X_i = \|\nabla U^*_{f,i} - \frac{1}{a}(\mathbf{\Sigma}_c \phi_i + \mathbf{\Sigma}_{f,i})\|^2_{\omega_i}, \\ Y_i = \sum_{K \in \mathcal{K}(\omega_i)} h^{-1} \|[U^*_{f,i} + U_{c,i}]\|^2_{\partial K \setminus \partial \omega_i}, \\ Z_i = \|h^{-1/2} U^*_{f,i}\|^2_{\partial \omega_i \setminus \Gamma}, \\ W_i = \|\frac{h}{a}(f\psi_i + \nabla \cdot (\mathbf{\Sigma}_c \phi_i + \mathbf{\Sigma}_{f,i}))\|^2_{\omega_i}. \end{cases} \tag{5.3}$$

We do not take the a dependent coefficient into account since it multiplies all expressions with the same factor. The four error indicators in equation (5.3) are easy and cheep to calculate.

Since we are not primarily interested in the absolute size of these estimators, just how they compare to each other, we do not use a tolerance as a stopping criteria. Below we will sketch a simple adaptive algorithm based on the four estimators in equation (5.3).

1. Calculate the solution the solution to equation (2.14) $\mathbf{\Sigma}$ using small patches with low resolution.

2. Calculate the four error indicators on each patch, equation (5.3).

3. For large values in X_i, Y_i, and, W_i refine the mesh once for patch i.

4. For large values in Z_i increase the patch size by one layer for patch i.

5. Stop if the solution is sufficiently good or go back to 1.

6 Numerical Examples

We will show numerical examples in 2D using the the permeability a of the tenth SPE comparative solution project for oil reservoir simulation. An often used method in oil drilling is to inject water into the oil reservoir in order to move the oil to a producer. Quick numerical simulation is important in this process in order to make correct decisions. The two phase flow in the oil reservoir is often modeled by a non linear transport equation. This equation is coupled with an elliptic equation for the pressure in the oil reservoir, see [2] for a review on this subject. The pressure gradient is then passed to the transport equation. The pressure equation basically takes the form of our model problem (1.1).

We take the permeability a from the lowest layer of the Upper Ness formation in all examples below, see Figure 1. The permeability is represented as a piecewise constant function on a grid consisting of 220×60 rectangular elements. We let $f = 1$ in the lower left corner simulating an injector and $f = -1$ in the upper right corner simulating a producer. In Figure 6 we plot reference solutions calculated on a 440×120 grid. In the examples we use a coarse rectangular mesh consisting of 55×15 blocks.

6.1 First Example: Dependency on the Number of Layers

In this first example we study how the method converge to a reference solution on the fine mesh when we use the same fine mesh for all patches and varies the patch size starting with one layer for all patches, then two layers for all, and so on. The quality of the reference solution is not so crucial here since we are, for the moment, only interested in the effect of the number of layers. We therefore use a reference solution calculated on 220×60 cells. We get a constant h on all patches by performing two uniform refinements. This means that if we would use maximum size patches, $\omega_i = \Omega$ we would get the same result from our method as we get from the reference solution. This is of course not interesting in practice but good to know when we interpret the

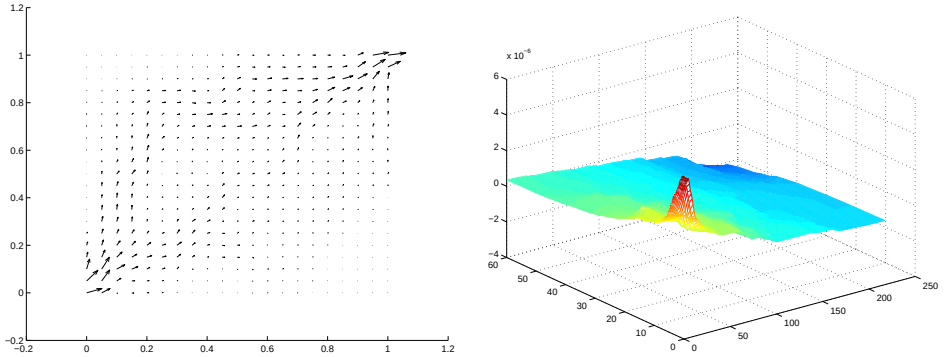


Figure 6: Above we see the reference solution, (left) flux $-\mathbf{\Sigma}$ and (right) pressure u.

results of our method using different sized patches. In this example we use no adaptivity, the same size on all patches. In Figure 7 we display the max norm error compared to the reference solution in log scale versus number of layers.

We note that we have exponential decay in the the error both in pressure and flux. Increasing the number of layers further will bring us closer and closer to the reference solution but already after two layers we get a very good agreement and taking into account the discretization error in the reference solution it is probably not of interest to use more than two layers in this example without refining the mesh.

6.2 Second Example: The Adaptive Algorithm

Next we will see how the adaptive algorithm works for the same problem. We use a reference solution calculated on a mesh that uses 440×120 elements and coarsened the data a two times to a 55×15 grid which means that the reference solution now has good quality. We start by using one layer patches with one refinement and then perform two iterations of the adaptive algorithm above. In each iteration we refine the mesh in 35% of the patches according to X, Y, and W in equation (5.3) and increase the number of layers in 35% of the patches according to Z in equation (5.3). In Figure 8 we give a graphic representation of the number of layers used for the local problems. In Figure 9 we show the number of refinements in each patch after two and three iterations.

We see clearly that more computational work is put in to certain areas, in particular close to the injector and the producer. We also compare the solutions to the reference solution calculates on the 440×120 grid. We

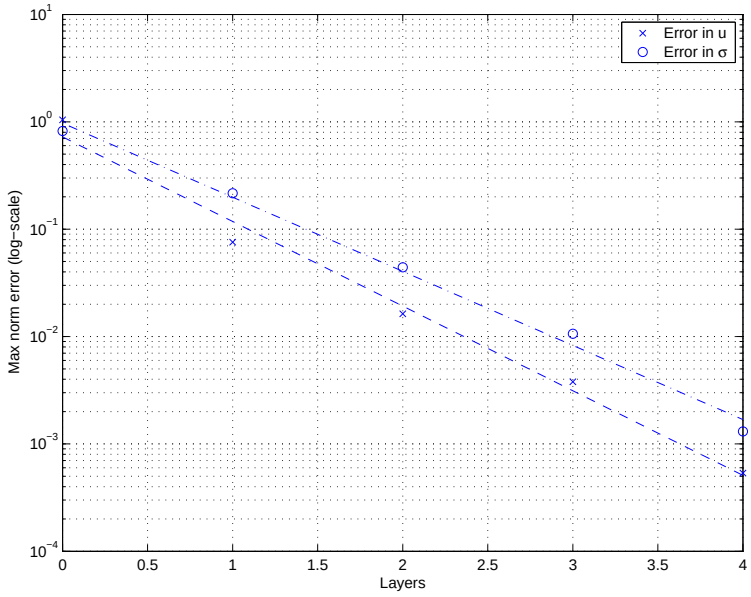


Figure 7: Max norm error (compared to reference solution) in log scale versus number of layers.

present the error in Figure 10.

In the upper left corner we have the standard Galerkin solution. The relative error compared to the reference solution in energy norm is 106% and we see clearly that the error has the same structure as the solution itself, compare $|\boldsymbol{\sigma}|$. In the upper right corner we see the error after one refinement and one layer patches. The error is now decreased to 16% but we still see a similar structure as the solution. After two refinements we solve more accurate fine scale problems close to the injector and the producer and we also see that the error decreases in these areas, the relative error in energy norm is now down to 10%. Finally after three iterations the error drops to 8% and we see from Figure 10, in the lower left corner, that the error is scattered over the domain. Remember that we would need three refinements for all patches and several layers to come very close to the reference solution. So the adaptive algorithm gives a substantial improvement at a low cost which was exactly our goal.

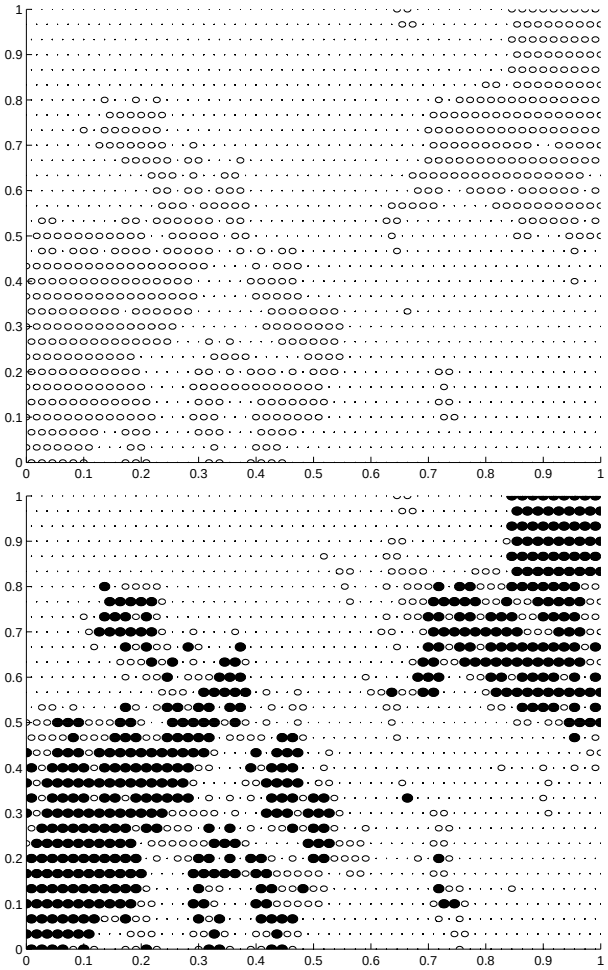


Figure 8: After the second iteration (above) we have increased 35% of the patches to two layers, these patches are indicated by circles. After the third iteration (below) again we increase 35% of the patches. This means that we have one layer patches on most patches illustrated by small dots, then we have some patches of two layers illustrated by medium circles and finally the three layer patches illustrated by the large dots.

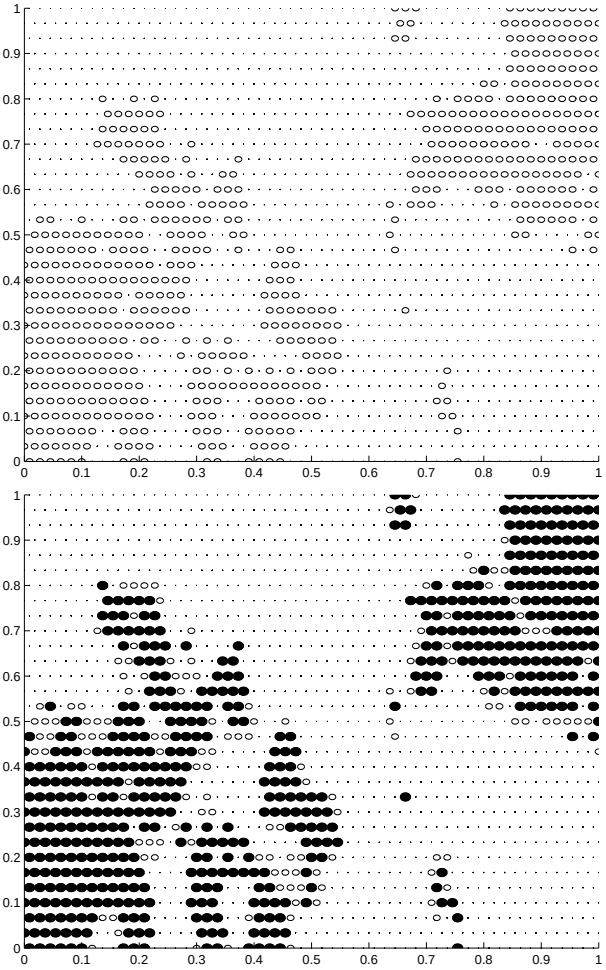


Figure 9: After the second iteration (above) we are using two refinements in 35% of the patches, these patches are indicated by a larger circles. After the third iteration (below) again we refine the mesh uniformly in 35% of the patches. This means that we have patches with one, two and three refinements represented by small dots, medium circles, and large dots.

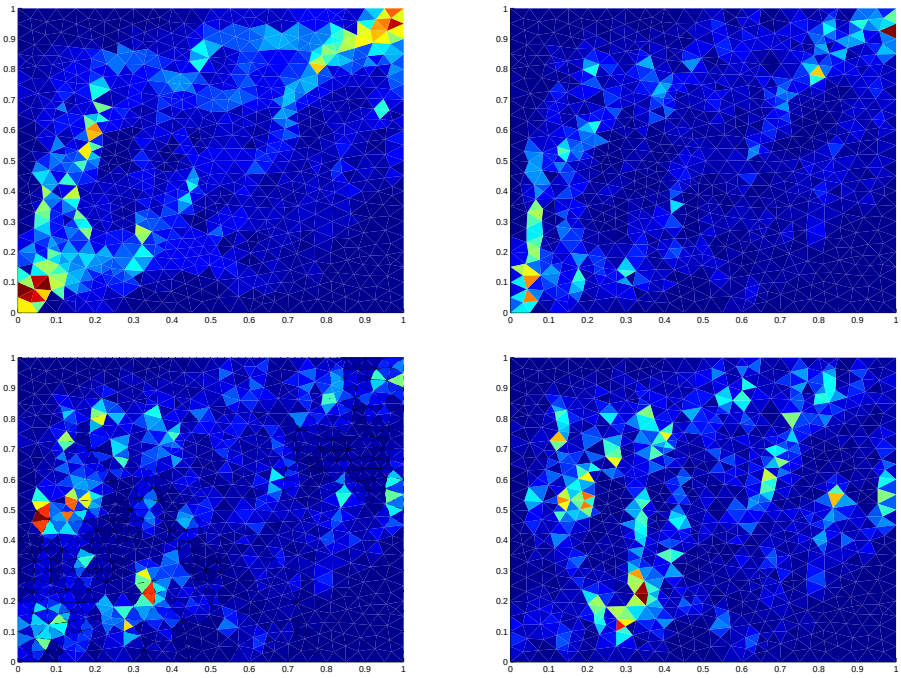


Figure 10: Upper left corner we have the relative error in energy norm for standard Galerkin (106%). Upper right corner is the error after using one refinement and one layer patches (16%). Lower left corner is the error after the second iteration (10%) and lower right is the error after three iterations (8%). We use rectangular elements in the computation but the plot function transforms this into triangles.

References

[1] J.E. Aarnes, K.-A. Lie, and A. B. Rustad, *Modelling of Multiscale Structures in Flow Simulations for Petroleum Reservoirs,* submitted.

[2] J.E. Aarnes, T. Gimse, and K.-A. Lie, *An introduction to the numerics of flow in porous media using Matlab,* Geometrical Modelling Numerical Simulation and Optimization, Industrial Mathematics at SINTEF. Springer Verlag.

[3] R. A. Adams, *Sobolev spaces,* volume 65 of Pure and Applied Mathematics, Academic Press, New York, 1975.

[4] T. Arbogast and S. L. Bryant, *A two-scale numerical subgrid technique for waterflood simulations,* SPE J., 2002.

[5] R. Becker, P. Hansbo, and M. G. Larson, *Energy norm estimation for discontinuous Galerkin methods,* Comput. Methods Appl. Mech. Engrg. 192, 723-733, 2003.

[6] W. Bangerth and R. Rannacher *Adaptive Finite Element Methods for Differential Equations* Birkhäuser, 2003.

[7] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods,* Springer Verlag, 1994.

[8] F. Brezzi, J. Douglar JR., and L. Marini, *Two families of mixed finite elements for second order elliptic problems,* Numer. Math., 47, 217-235, 1985.

[9] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, *Computational differential equations,* Studentlitteratur, 1996.

[10] V. Girault, P.-A. Raviart, *Finite element approximation of the Navier-Stokes equation,* Springer Verlag, Berlin, 1979.

[11] P. Hood and C. Taylor, *A numerical solution of the Navier-Stokes equations using the finite element techniques,* Comp. and Fluids, 1, 73-100, 1973.

[12] T. Y. Hou and X.-H. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media,* J. Comput. Phys. 134 (1997) 169-189.

[13] T. J.R. Hughes, *Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and*

the origins of stabilized methods, Comput. Methods Appl. Mech. Engrg. 127 (1995) 387-401.

[14] T. J.R. Hughes, G. R. Feijóo, L. Mazzei and Jean-Baptiste Quincy, *The variational multiscale method - a paradigm for computational mechanics,* Comput. Methods Appl. Mech. Engrg. 166 (1998) 3-24.

[15] M. G. Larson and A. Målqvist, *Adaptive variational multiscale methods based on a posteriori error estimation: Duality techniques for elliptic problems,* Lecture notes in: Computational Science and Engineering. B. Enquist, P. Löstedt, O. Runborg (eds.) Vol. 44, 181-193, 2005.

[16] M. G. Larson and A. Målqvist, *Adaptive variational multiscale methods based on a posteriori error estimation,* Proceedings of the 4th European Congress on Computational Methods in Applied Sciences and Engineering, P. Neittaanmäki, T. Rossi, S. Korotov, E. Onate, J. Periaux, and D. Knörzer (eds.), 2004.

[17] M. G. Larson and A. Målqvist, *Adaptive Variational Multiscale Methods Based on A Posteriori Error Estimation: Energy Norm Estimates for Elliptic Problems,* submitted.

[18] C. Lovadina and R. Stenberg, *Energy norm a posteriori error estimates for mixed finite element methods,* Helsinki University of Technology. Research Report A473. September 2004.

[19] A. Målqvist and M. G. Larson, *A Posteriori Error Estimates in the Energy Norm for Mixed Finite Element Methods,* submitted.

[20] J.-C. Nédélec, *A new family of mixed finite elements in* $\mathbf{R}^3$*,* Numer. Math., 50, 57-81, 1986.

[21] P. Raviart and J. Thomas, *A mixed finite element method for second order elliptic problems,* in Mathematical Aspects of the Finite Element Method. Lecture Notes in Math. 606, Springer-Verlag, 292-315, 1977.

Goal Oriented Adaptivity for Coupled Flow and Transport Problems with Applications in Oil Reservoir Simulation.

Axel Målqvist* and Mats G. Larson†

August 19, 2005

Abstract

Goal oriented adaptivity has become an important and widely used technique for solving partial differential equations at a reduced computational cost. The technique is based on a computable estimate of the error in a linear functional of interest and an adaptive strategy that improves the solution, for instance by refining the mesh locally according to the error estimator.

In this paper we extend this theory to a multi-physical setting involving a coupled set of equations. Here adaptive algorithms become even more crucial since we do not only need estimators that indicates in which part of the domain the solution needs to be improved but also which equation that contributes most to the error and thus needs to be solved more accurately.

We derive a goal oriented error estimator for a system of coupled pressure and transport that serves as a basic model for oil reservoir simulation. We present an extensive numerical example where we explain in detail how the adaptive algorithm works in practice.

1 Introduction

In this paper we develop adaptive algorithms based on goal oriented a posteriori error estimates for a coupled pressure-transport problem, which serves as a basic model problem for oil reservoir simulation. The pressure equation determines the convection field which is used in the transport problem. The error estimate is based on duality arguments and identifies the individual

*Graduate Research Assistant, Department of Mathematics, Chalmers University of Technology, Göteborg, axel@math.chalmers.se.

†Professor of Applied Mathematics, Department of Mathematics, Umeå University, Umeå, mats.larson@math.umu.se.

effects of discretization errors in the pressure and transport solvers on the quantity of interest. The approach is fairly general and may be employed to other multi-physics problems, for instance fluid structure interaction, and thermal expansion problems in solid mechanics.

The Model Problem. We consider the following transport problem,

$$\begin{cases} \dot{c} + \nabla \cdot (\boldsymbol{\sigma} c) - \epsilon \triangle c = g \quad \text{in } \Omega \times (0, T], \\ \boldsymbol{n} \cdot \nabla c = 0 \quad \text{on } \Gamma, \\ c = c_0 \quad \text{for } t = 0, \end{cases} \tag{1.1}$$

where Ω is a polygonal domain in $\mathbf{R}^n$ with boundary Γ, $g \in L^2(\Omega)$ is given data, $c_0 \in L^2(\Omega)$ is initial data, and the flux $\boldsymbol{\sigma}$ is determined by the following equation,

$$\begin{cases} \frac{1}{a}\boldsymbol{\sigma} - \nabla u = 0 \quad \text{in } \Omega, \\ -\nabla \cdot \boldsymbol{\sigma} = f \quad \text{in } \Omega, \\ \boldsymbol{n} \cdot \boldsymbol{\sigma} = 0 \quad \text{on } \Gamma. \end{cases} \tag{1.2}$$

Here $a \in L^\infty(\Omega)$ is positive and $f \in L^2(\Omega)$. Assuming that $\int_\Omega f \, dx = 0$ we have a well posed problem with a solution $u \in H^1(\Omega)/\mathbf{R}$ and $\boldsymbol{\sigma} \in \boldsymbol{V} = \{\boldsymbol{v} \in H(\mathrm{div}; \Omega) : \boldsymbol{n} \cdot \boldsymbol{v} = 0 \text{ on } \Gamma\}$. See [10] for definitions of these function spaces.

Previous Work. There have been many works on goal oriented adaptivity during the last decade, see for instance Johnson et.al. [12, 15, 13], Rannacher et.al. [6], and the references therein. The treatment of single-physics problems is fairly well understood, while the theory has not yet been applied to multi-physics problems to any large extent. Coupled flow and transport problems appear in numerous applications, for instance in simulation of oil reservoirs, see Aarnes and Lie et.al. [2, 1], Wheeler et.al. [11], Arbogast et.al. [4], for background on this topic.

New Contributions. In this paper we develop adaptive algorithms for a coupled pressure-transport equation, motivated by oil reservoir simulation, based on a posteriori error estimates. Often one is particularly interested in the production of oil in a reservoir simulation. The production may be expressed as a linear functional (the localized value of the solution at the producer) of the solution and thus the problem is ideal for a goal oriented approach. We also have a strong multiscale scale structure in the permeability that requires different degrees of refinement in different parts of the domain. A third reason and perhaps the most important is that this problem consists of several partial differential equations that most likely need different meshes and the choice of suitable meshes is difficult to do by hand.

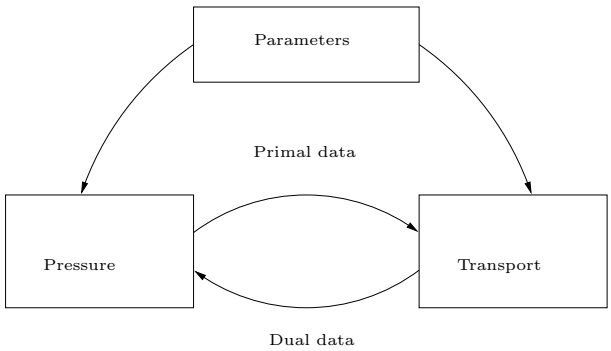


Figure 1: Information flow between solvers and data base.

When one must determine a large number of discretization parameters such as mesh sizes, time steps, and so on automatic tuning of these parameters by an adaptive algorithm becomes crucial. The need for automatic tuning of parameters thus becomes more and more apparent when several solvers are connected in a multi-physics simulation.

We present a duality based a posteriori error estimate for a linear functional of the error together with an adaptive algorithm and numerical examples on a simple model problem for oil reservoir simulation. To illustrate the flow of data in the adaptive algorithm we present a schematic picture in Figure 1. The box with parameters symbolizes the known data in the equations such as g, c_0, Ω, a, f and so on. The box to the left is the solver of the pressure equation that feeds the transport solver (to the right) with the calculated flux $\boldsymbol{\sigma}$. In the transport solver we solve the primal transport problem but also a dual transport problem driven by a given weight that determines in what sense we control the error (the type of linear functional). The primal and dual transport solutions are then combined according to the error estimates and used as data to a dual pressure solver. At this point we are ready to refine the meshes according to the error estimates and start over again by solving the primal pressure problem.

Outline. The remainder of the paper is organized as follows: in Section 2 we present the finite element formulation of the model problem; in Section 3 we derive error estimates that leads to an adaptive algorithm in Section 4; and finally in section 5 we present extensive numerical examples.

2 The Finite Element Method

We introduce some notations before we can present a discrete version of the weak form of equation (1.1). We let,

$$(v, w) = \int_\Omega v \cdot w \, dx, \tag{2.1}$$

for all $v, w \in L^2(\Omega)$ and $\|v\|_\Omega$ be the $L^2(\Omega)$ norm of v. We let V_h be the space of continuous piecewise linear functions on a triangulation $\mathcal{K} = \{K\}$, which satisfies the minimal angle condition, of Ω. The mesh parameter h is a piecewise constant function such that on each element K, $h = \text{diam}(K)$. Furthermore, we let

$$\pi : L^2(\Omega) \to V_h, \tag{2.2}$$

be the Scott-Zhang interpolant onto the space of piecewise linears V_h, see [9]. We also recall for later use that π is stable in $H^1(\Omega)$.

2.1 The Transport Problem

Since $\boldsymbol{\sigma}$ is an unknown solution to equation (1.2) we assume, for the moment, that we have a discrete approximation $\mathbf{\Sigma}$ such that $\boldsymbol{n} \cdot \mathbf{\Sigma} = 0$ on Γ. We will use the cG1-cG1 method to approximate the transport equation, see [13]. This means that $C : (0, T] \to V_h$ will be the solution to:

$$\int_{I_k} (\dot{C}, v) + (\nabla \cdot (\mathbf{\Sigma} C), v) + (\epsilon \nabla C, \nabla v) \, dt = \int_{I_k} (g, v) \, dt, \quad \text{for all } v \in V_h, \tag{2.3}$$

where $k = 1, ..., n$, $I_k = [t_{k-1}, t_k]$, $t_{k-1} < t_k$, $t_0 = 0$, and $t_n = T$. We also introduce the notation $\triangle t_k = t_k - t_{k-1}$. We define $\triangle t$ as a piecewise constant function in time. For $t \in I_k$, $\triangle t = \triangle t_k$.

The continuous piecewise linear basis functions we use in time gives the following identities,

$$\int_{I_k} (\dot{C}, v) \, dt = (C(t_k) - C(t_{k-1}), v), \tag{2.4}$$

$$\int_{I_k} (\nabla \cdot (\mathbf{\Sigma} C), v) \, dt = \triangle t_k \cdot (\nabla \cdot (\mathbf{\Sigma} (\frac{C(t_k) + C(t_{k-1})}{2})), v), \tag{2.5}$$

$$\int_{I_k} (\epsilon \nabla C, \nabla v) \, dt = \triangle t_k \cdot (\epsilon \nabla \frac{C(t_k) + C(t_{k-1})}{2}, \nabla v), \tag{2.6}$$

and

$$\int_{I_k} (g, v) \, dt = \triangle t_k \cdot (g, v), \tag{2.7}$$

since we assume g to be independent of time. We introduce the notation $C(t_k) = C_k$ and get the following finite element method: find $C_k \in V_h$ such that

$$\begin{aligned}(C_k + \frac{\triangle t_k}{2}\nabla\cdot(\mathbf{\Sigma}C_k), v) + \frac{\triangle t_k}{2}(\epsilon\nabla C_k, \nabla v) \\ = (C_{k-1} - \frac{\triangle t_k}{2}\nabla\cdot(\mathbf{\Sigma}C_{k-1}), v) \\ - \frac{\triangle t_k}{2}(\epsilon\nabla C_{k-1}, \nabla v) + \triangle t_k(g, v),\end{aligned} \tag{2.8}$$

for all $v \in V_h$, $k = 1, ..., n$, and C_0 equal to the $L^2(\Omega)$ projection of c_0 onto V_h.

To stabilize the cG1-cG1 method for convection dominated problems, ϵ small compared to $\mathbf{\Sigma}$, we use the streamline diffusion method. We replace the test function v by $v + \delta\nabla\cdot(\mathbf{\Sigma}v)$ in two terms of equation (2.3), the first one is the convection term $(\nabla\cdot(\mathbf{\Sigma}C), v)$ and the second one is the load term (f, v). The parameter δ is defined as a piecewise constant on $\mathcal{K}$

$$\delta = \frac{D}{\sqrt{(\frac{1}{\triangle t^2} + \frac{|\bar{\mathbf{\Sigma}}|^2}{h^2})}}, \tag{2.9}$$

D will throughout this paper constants and $\bar{\mathbf{\Sigma}}$ is the mean value of $\mathbf{\Sigma}$ on the current element. For more information on streamline-diffusion see [14, 12, 13].

2.2 The Flow Problem

We let $\mathcal{Q} = \{K\}$ be a partition, possibly different from $\mathcal{K}$, of Ω, $\boldsymbol{V}_h$ be the lowest order Raviart-Thomas space [20, 19] for the flux, and W_h be the space of piecewise constants on $\mathcal{Q}$. The finite element method then reads: find $\mathbf{\Sigma} \in \boldsymbol{V}_h$ and $U \in W_h$ such that,

$$\begin{cases} (\frac{1}{a}\mathbf{\Sigma}, \boldsymbol{v}) + (U, \nabla\cdot\boldsymbol{v}) = 0, \\ -(\nabla\cdot\mathbf{\Sigma}, w) = (f, w), \end{cases} \tag{2.10}$$

for all $\boldsymbol{v} \in \boldsymbol{V}_h$ and $w \in W_h$.

3 A Posteriori Error Estimation

We do not focus our attention on tracking constants in this paper, instead we seek to derive error estimates that scale properly with the error modulo constants. Constants denoted D does not depend on the mesh size h. We

will not consider the streamline-diffusion term in the error analysis. For an extensive discussion on streamline-diffusion method we guide the reader to [14, 15].

3.1 Error Equation

We get the following Galerkin orthogonality for the error $e = c - C$ by subtracting equation (2.3) from the weak form of equation (1.1),

$$\int_{I_k} (\dot{e}, v) + (\nabla \cdot (\boldsymbol{\Sigma} e), v) + (\epsilon \nabla e, \nabla v)\, dt = -\int_{I_k} (\nabla \cdot ((\boldsymbol{\sigma} - \boldsymbol{\Sigma})c), v)\, dt \quad (3.1)$$

$$= \int_{I_k} ((\boldsymbol{\sigma} - \boldsymbol{\Sigma}), c\nabla v)\, dt, \quad (3.2)$$

for all $v \in \boldsymbol{V}_h$. In the last equality we used that $\boldsymbol{n} \cdot (\boldsymbol{\sigma} - \boldsymbol{\Sigma})$ vanishes on the boundary Γ.

3.2 Dual Transport Problem

Next we introduce the dual problem,

$$\begin{cases} -\dot{\phi} - \boldsymbol{\Sigma} \cdot \nabla\phi - \epsilon\triangle\phi = \psi \quad \text{in } \Omega \times (0, T], \\ \boldsymbol{n} \cdot \nabla\phi = 0 \quad \text{on } \Gamma, \\ \phi = 0 \quad \text{for } t = T, \end{cases} \quad (3.3)$$

where $\psi \in H^{-1}(\Omega)$. Note that this dual problem is solved backwards in time starting with initial data at time T.

3.3 Error Estimate of the Transport Problem

Proposition 3.1 *For an arbitrary distribution $\psi \in H^{-1}(\Omega)$ it holds,*

$$\int_0^T (e, \psi)\, dt \le \int_0^T D \sum_{K \in \mathcal{K}} \rho_K(C)(\triangle t \|\dot{\phi}\|_K + h\|\nabla\phi\|_K)\, dt \quad (3.4)$$

$$+ |(\boldsymbol{\sigma} - \boldsymbol{\Sigma}, \int_0^T c\nabla\pi\bar{\phi}\, dt)|,$$

where,

$$\bar{\phi} = \frac{1}{\triangle t_k} \int_{I_k} \phi\, dt, \quad (3.5)$$

for $t \in I_k$ and

$$\rho_K(C) = \|\dot{C} + \nabla \cdot (\boldsymbol{\Sigma} C) - \epsilon\triangle C - g\|_K + h^{-1/2}\|\epsilon[\boldsymbol{n} \cdot \nabla C]\|_{\partial K}. \quad (3.6)$$

Proof. Multiplying the dual problem (3.3) by the error and integrating by parts, in time for the first term and in space for the second and third term, where boundary terms vanish due to the boundary conditions, we get the following identities

$$\int_0^T (e,\psi)\,dt = \int_0^T (e, -\dot{\phi} - \boldsymbol{\Sigma}\cdot\nabla\phi - \epsilon\triangle\phi)\,dt \tag{3.7}$$

$$= \int_0^T (\dot{e},\phi) - (\boldsymbol{\Sigma}e,\nabla\phi) + (\epsilon\nabla e,\nabla\phi)\,dt \tag{3.8}$$

$$= \int_0^T (\dot{e},\phi) + (\nabla\cdot(\boldsymbol{\Sigma}e),\phi) + (\epsilon\nabla e,\nabla\phi)\,dt. \tag{3.9}$$

We define the time mean value of ϕ in each time slab,

$$\bar{\phi} = \frac{1}{\triangle t_k}\int_{I_k}\phi\,dt, \tag{3.10}$$

for $t \in I_k$, and use equation (3.1) to subtract $\pi\bar{\phi} \in V_h$, from equation (3.7),

$$\int_0^T (e,\psi)\,dt = \sum_{k=1}^n \int_{I_k} (\dot{e},\phi) + (\nabla\cdot(\boldsymbol{\Sigma}e),\phi) + (\epsilon\nabla e,\nabla\phi)\,dt \tag{3.11}$$

$$= \sum_{k=1}^n \int_{I_k} (\dot{e},\phi-\pi\bar{\phi}) + (\nabla\cdot(\boldsymbol{\Sigma}e),\phi-\pi\bar{\phi})\,dt \tag{3.12}$$

$$+ \sum_{k=1}^n \int_{I_k} (\epsilon\nabla e,\nabla(\phi-\pi\bar{\phi})) + (\boldsymbol{\sigma}-\boldsymbol{\Sigma}, c\nabla\pi\bar{\phi})\,dt.$$

Next we use the fact that c is the exact solution to (1.1) to eliminate c and we also recall that $\boldsymbol{\sigma}$ and $\boldsymbol{\Sigma}$ are independent of time to simplify the last

term on the right hand side in (3.12).

$$\int_0^T (e,\psi)\,dt = \int_0^T (g - \dot{C} - \nabla\cdot(\mathbf{\Sigma}C), \phi - \pi\bar{\phi}) + (\epsilon\nabla C, \nabla(\phi - \pi\bar{\phi}))\,dt \tag{3.13}$$

$$+ (\boldsymbol{\sigma} - \mathbf{\Sigma}, \int_0^T c\nabla\pi\bar{\phi}\,dt)$$

$$\leq \int_0^T \sum_{K\in\mathcal{K}} (g - \dot{C} - \nabla\cdot(\mathbf{\Sigma}C), \phi - \pi\bar{\phi})_K\,dt \tag{3.14}$$

$$+ \int_0^T \sum_{K\in\mathcal{K}} (\epsilon\nabla C, \nabla(\phi - \pi\bar{\phi}))_K\,dt + (\boldsymbol{\sigma} - \mathbf{\Sigma}, \int_0^T c\nabla\pi\bar{\phi}\,dt)$$

$$\leq \int_0^T \sum_{K\in\mathcal{K}} \|g - \dot{C} - \nabla\cdot(\mathbf{\Sigma}C) + \epsilon\triangle C\|_K \|\phi - \pi\bar{\phi}\|_K\,dt \tag{3.15}$$

$$+ \int_0^T \sum_{K\in\mathcal{K}} |\frac{1}{2}(\epsilon[\boldsymbol{n}\cdot\nabla C], \phi - \pi\bar{\phi})_{\partial K}|\,dt$$

$$+ (\boldsymbol{\sigma} - \mathbf{\Sigma}, \int_0^T c\nabla\pi\bar{\phi}\,dt),$$

where the jumps $[\cdot]$ in equation (3.15) denotes the difference in function value over an interior face in the mesh. These jumps appears since C does not have continuous derivative over element edges, see for instance [13, 16, 6]. Using standard estimates, see for instance [16], we can estimate the jump term in the following way,

$$(\epsilon[\boldsymbol{n}\cdot\nabla C], \phi - \pi\bar{\phi})_{\partial K} \leq h^{-1/2}\|\epsilon[\boldsymbol{n}\cdot\nabla C]\|_{\partial K} h^{1/2}\|\phi - \pi\bar{\phi}\|_{\partial K} \tag{3.16}$$

$$\leq Dh^{-1/2}\|\epsilon[\boldsymbol{n}\cdot\nabla C]\|_{\partial K}\|\phi - \pi\bar{\phi}\|_K \tag{3.17}$$

$$+ Dh^{-1/2}\|\epsilon[\boldsymbol{n}\cdot\nabla C]\|_{\partial K} h\|\nabla(\phi - \pi\bar{\phi})\|_K).$$

Since π is stable in $H^1(\Omega)$ we have $\|\nabla(\phi - \pi\bar{\phi})\|_K \leq D(\|\nabla\phi\|_K + \|\nabla\bar{\phi}\|_K)$. We further assume $\|\nabla\bar{\phi}\| \leq D\|\nabla\phi\|$. We conclude,

$$\begin{aligned}\int_0^T (e,\psi)\,dt \leq \int_0^T \sum_{K\in\mathcal{K}} \|g - \dot{C} - \nabla\cdot(\boldsymbol{\Sigma}C) + \epsilon\triangle C\|_K \|\phi - \pi\bar{\phi}\|_K\,dt \quad (3.18)\\ + D\int_0^T \sum_{K\in\mathcal{K}} h^{-1/2}\|\epsilon[\boldsymbol{n}\cdot\nabla C]\|_{\partial K}\|\phi - \pi\bar{\phi}\|_K\,dt\\ + D\int_0^T \sum_{K\in\mathcal{K}} h^{-1/2}\|\epsilon[\boldsymbol{n}\cdot\nabla C]\|_{\partial K} h\|\nabla\phi\|_K\,dt\\ + (\boldsymbol{\sigma} - \boldsymbol{\Sigma}, \int_0^T c\nabla\pi\bar{\phi}\,dt).\end{aligned}$$

Lets consider the interpolation term $\|\phi - \pi\bar{\phi}\|_K$ on a given time interval I_k. We note that,

$$\|\phi - \pi\bar{\phi}\|_K \leq \|\phi - \bar{\phi}\|_K + \|\bar{\phi} - \pi\bar{\phi}\|_K, \quad (3.19)$$

but since $\bar{\phi}$ is a mean value in time the first term can be estimated by $\|\phi - \bar{\phi}\|_K \leq D\triangle t_k\|\dot{\phi}\|_K$ and for the second term we have interpolation which gives, $\|\bar{\phi} - \pi\bar{\phi}\|_K \leq Dh\|\nabla\phi\|_K$ i.e.

$$\|\phi - \pi\bar{\phi}\|_K \leq D(h\|\nabla\phi\|_K + \triangle t_k\|\dot{\phi}\|_K), \quad (3.20)$$

if we use the result from equation (3.20) in equation (3.18) we end up with,

$$\begin{aligned}\int_0^T (e,\psi)\,dt \leq D\int_0^T \sum_{K\in\mathcal{K}} \|g - \dot{C} - \nabla\cdot(\boldsymbol{\Sigma}C) + \epsilon\triangle C\|_K h\|\nabla\phi\|_K\,dt \quad (3.21)\\ + D\int_0^T \sum_{K\in\mathcal{K}} \|g - \dot{C} - \nabla\cdot(\boldsymbol{\Sigma}C) + \epsilon\triangle C\|_K \triangle t\|\dot{\phi}\|_K\,dt\\ + D\int_0^T \sum_{K\in\mathcal{K}} h^{-1/2}\|\epsilon[\boldsymbol{n}\cdot\nabla C]\|_{\partial K}(h\|\nabla\phi\|_K + k\|\dot{\phi}\|_K)\,dt\\ + (\boldsymbol{\sigma} - \boldsymbol{\Sigma}, \int_0^T c\nabla\pi\bar{\phi}\,dt).\end{aligned}$$

We now introduce the following element indicator function,

$$\rho_K(C) = \|\dot{C} + \nabla\cdot(\boldsymbol{\Sigma}C) - \epsilon\triangle C - g\|_K + h^{-1/2}\|\epsilon[\boldsymbol{n}\cdot\nabla C]\|_{\partial K}. \quad (3.22)$$

This proves the proposition. □

The last term in the estimate accounts for the influence of the error in the flux $\boldsymbol{\sigma} - \boldsymbol{\Sigma}$ and may be interpreted as a modeling error term. Clearly

$(\boldsymbol{\sigma}-\boldsymbol{\Sigma},\boldsymbol{\omega})$ with $\boldsymbol{\omega}=\int_0^T c\nabla\pi\bar{\phi}\,dt$ is the functional of the flux which needs to be controlled to guarantee overall accuracy in the transport computation. We are thus naturally led to deriving a goal oriented error estimate for this particular functional.

3.4 Dual Elliptic Problem

To handle the last term in Proposition 3.1 we introduce another dual problem:

$$\begin{cases} \frac{1}{a}\boldsymbol{\chi}-\nabla\eta=\boldsymbol{\omega} & \text{in } \Omega,\\ -\nabla\cdot\boldsymbol{\chi}=0 & \text{in } \Omega,\\ \boldsymbol{n}\cdot\boldsymbol{\chi}=0 & \text{on } \Gamma, \end{cases} \tag{3.23}$$

and formulate the following proposition.

3.5 Error Estimate for the Mixed Problem

Proposition 3.2 *It holds,*

$$\begin{aligned} |(\boldsymbol{\sigma}-\boldsymbol{\Sigma},\boldsymbol{\omega})| \le D\sum_{K\in\mathcal{Q}}(\|\nabla U^*-\frac{1}{a}\boldsymbol{\Sigma}\|_K+h^{-1/2}\|[U^*]\|_{\partial K\setminus\Gamma})\|\boldsymbol{\chi}\|_K \\ +D\sum_{K\in\mathcal{Q}}h\|\nabla\cdot\boldsymbol{\Sigma}+f\|_K\|\nabla\eta\|_K, \end{aligned} \tag{3.24}$$

for arbitrary $U^*\in\bigoplus_{K\in\mathcal{Q}}P^1(K)$ *where* $P^1(K)$ *are piecewise linear polynomials on* K*.*

Proof. We have,

$$(\boldsymbol{\sigma}-\boldsymbol{\Sigma},\boldsymbol{\omega})=(\frac{1}{a}(\boldsymbol{\sigma}-\boldsymbol{\Sigma}),\boldsymbol{\chi})-(\boldsymbol{\sigma}-\boldsymbol{\Sigma},\nabla\eta) \tag{3.25}$$

$$=-(u,\nabla\cdot\boldsymbol{\chi})-(\frac{1}{a}\boldsymbol{\Sigma},\boldsymbol{\chi})+(\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}),\eta) \tag{3.26}$$

$$=-(\frac{1}{a}\boldsymbol{\Sigma},\boldsymbol{\chi})+(\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma}),\eta). \tag{3.27}$$

We let $P_K: W\to W_h$ be the elementwise L^2-projection and note, by using the second part of equation (2.10), that we can add $(f+\nabla\cdot\boldsymbol{\Sigma},P_K\eta)$ to equation (3.27). We also introduce the arbitrary function U^* by adding,

$$-(U^*,\nabla\cdot\boldsymbol{\chi})=0. \tag{3.28}$$

This gives,

$$(\boldsymbol{\sigma} - \boldsymbol{\Sigma}, \boldsymbol{\omega}) = -(U^*, \nabla \cdot \boldsymbol{\chi}) - (\frac{1}{a}\boldsymbol{\Sigma}, \boldsymbol{\chi}) - (\nabla \cdot \boldsymbol{\Sigma} + f, \eta - P_K\eta) \tag{3.29}$$

$$= \sum_{K \in \mathcal{Q}} (\nabla U^* - \frac{1}{a}\boldsymbol{\Sigma}, \boldsymbol{\chi})_K - (U^*, n \cdot \boldsymbol{\chi})_{\partial K} \tag{3.30}$$

$$- (\nabla \cdot \boldsymbol{\Sigma} + f, \eta - P_K\eta).$$

Since we aim to get an element error indicator we would like to use the Cauchy-Schwartz inequality. This is simple for the first and third term in equation (3.30) but a bit more complicated for the second term.

We use the proof of Theorem 3.1 in [18] to get the following bound,

$$(U^*, n \cdot \boldsymbol{\chi})_{\partial K} \leq Dh^{-1/2} \|[U^*]\|_{\partial K \setminus \Gamma} (\|\boldsymbol{\chi}\|_K + h\|\nabla \cdot \boldsymbol{\chi}\|_K) \tag{3.31}$$

$$= Dh^{-1/2} \|[U^*]\|_{\partial K \setminus \Gamma} \|\boldsymbol{\chi}\|_K,$$

since $\nabla \cdot \boldsymbol{\chi} = 0$ and $\boldsymbol{n} \cdot \boldsymbol{\chi} = 0$ on Γ. Using equation (3.31) and Cauchy-Schwartz inequality in equation (3.30). We end up with,

$$(\boldsymbol{\sigma} - \boldsymbol{\Sigma}, \boldsymbol{\omega}) = \sum_{K \in \mathcal{K}} (\nabla U^* - \frac{1}{a}\boldsymbol{\Sigma}, \boldsymbol{\chi})_K - (U^*, n \cdot \boldsymbol{\chi})_{\partial K} \tag{3.32}$$

$$- (\nabla \cdot \boldsymbol{\Sigma} + f, \eta - P_K\eta)$$

$$\leq D \sum_{K \in \mathcal{K}} (\|\nabla U^* - \frac{1}{a}\boldsymbol{\Sigma}\|_K + h^{-1/2} \|[U^*]\|_{\partial K \setminus \Gamma}) \|\boldsymbol{\chi}\|_K \tag{3.33}$$

$$+ D \sum_{K \in \mathcal{K}} h\|\nabla \cdot \boldsymbol{\Sigma} + f\|_K \|\frac{1}{h}(\eta - P_K\eta)\|_K \tag{3.34}$$

$$\leq D \sum_{K \in \mathcal{K}} (\|\nabla U^* - \frac{1}{a}\boldsymbol{\Sigma}\|_K + h^{-1/2} \|[U^*]\|_{\partial K \setminus \Gamma}) \|\boldsymbol{\chi}\|_K \tag{3.35}$$

$$+ D \sum_{K \in \mathcal{K}} h\|\nabla \cdot \boldsymbol{\Sigma} + f\|_K \|\nabla \eta\|_K,$$

using the interpolation estimate $\|(\eta - P_K\eta)/h\|_K \leq D\|\nabla\eta\|_K$. The proposition follows immediately. □

In Proposition 3.2 we introduce the function U^* instead of simply using U. It is known from [5, 8, 17, 21] that in order to get an optimal order error estimate when using a richer space for the flux then for the pressure, U needs to be improved. There are many ways to do this postprocessing and in this paper we will use the method formulated by Stenberg et. al. in [17, 21]. In the case of lowest order Raviart-Thomas base functions together with piecewise constants the idea is simply to derive a piecewise linear approximation

of U with the same mean value as U but with a gradient approximately equal to $\mathbf{\Sigma}/a$.

A short description of the method is to be found in [18]. We also present it here for completeness. We let $P^1(K)$ denote the space to piecewise linear functions on K.

Definition 3.1 *(Postprocessing method) Find U^* such that $U^*|_K = U_K^* \in P^1(K)$ where U_K^* is defined by*

$$P_K U_K^* = U_K, \tag{3.36}$$

and

$$(\nabla U^*, \nabla v)_K = (\frac{1}{a}\mathbf{\Sigma}, \nabla v)_K \quad \textit{for all } v \in (I - P_K)P^1(K). \tag{3.37}$$

With this definition together with the two propositions we are ready to present the main theorem of this paper.

Theorem 3.1 *It holds,*

$$\int_0^T (e, \psi)\, dt \leq D \sum_{K\in\mathcal{K}} \int_0^T \rho_K(C)(k\|\dot{\phi}\|_K + h\|\nabla\phi\|_K)\, dt \tag{3.38}$$

$$+ D \sum_{K\in\mathcal{Q}} (\|\nabla U^* - \frac{1}{a}\mathbf{\Sigma}\|_K + h^{-1/2}\|[U^*]\|_{\partial K\setminus\Gamma})\|\boldsymbol{\chi}\|_K \tag{3.39}$$

$$+ D \sum_{K\in\mathcal{Q}} h\|\nabla\cdot\mathbf{\Sigma} + f\|_K\|\nabla\eta\|_K, \tag{3.40}$$

where,

$$\bar{\phi} = \frac{1}{\triangle t_k}\int_{I_k} \phi\, dt, \tag{3.41}$$

for $t \in I_K$, U^* is defined in Definition 3.1 and

$$\rho_K(C) = \|\dot{C} + \nabla\cdot(\mathbf{\Sigma}C) - \epsilon\triangle C - g\|_K + h^{-1/2}\|\epsilon[\boldsymbol{n}\cdot\nabla C]\|_{\partial K}. \tag{3.42}$$

Proof. We simply combine the results in Proposition 3.1 and 3.2. □

We are now ready to formulate an adaptive algorithm based on this estimate. The parameters we want to tune are the mesh sizes of the transport problem and the elliptic problem. Remember that we have the possibility to choose different meshes for the two equations.

4 Adaptive Strategy

We construct an algorithm that improves a certain linear functional of the solution, $\int_0^T (c, \psi)\, dx$, ψ is chosen arbitrarily.

1. Calculate the solutions $\mathbf{\Sigma}$ and U to the elliptic problem (2.10) on a given initial mesh, $\mathcal{Q}$.
2. Calculate the solution to the transport problem C by solving equation (2.8) on an other initial mesh, $\mathcal{K}$.
3. Calculate an approximate solution of the dual transport problem (3.3), Φ, given ψ, on the same mesh as the transport problem or a uniformly refined version.
4. Calculate an approximation to $\boldsymbol{\omega} = \int_0^T c\nabla\pi\bar{\phi}\, dx$ from C and Φ.
5. Calculate the approximate solutions to the dual elliptic problem $\boldsymbol{\chi}$ and η on the mesh $\mathcal{Q}$ or perhaps a uniformly refined version.
6. Calculate the elementwise error indicators using the calculated primal and dual solutions,

$$\begin{cases} I_1^K = \int_0^T \rho_K(C)(k\|\dot{\phi}\|_K + h\|\nabla\phi\|_K)\, dt, \\ I_2^K = (\|\nabla U^* - \frac{1}{a}\mathbf{\Sigma}\|_K + h^{-1/2}\|[U^*]\|_{\partial K\setminus\Gamma})\|\boldsymbol{\chi}\|_K \\ \qquad\qquad +h\|\nabla\cdot\mathbf{\Sigma} + f\|_K\|\nabla\eta\|_K. \end{cases} \tag{4.1}$$

7. Derive the sums $I_1 = \sum_{K\in\mathcal{K}} I_1^K$ and $I_2 = \sum_{K\in\mathcal{Q}} I_2^K$.
8. If the indicators are sufficiently small we stop. If not we refine the meshes in the following way.
9. If $I_1 > EI_2$ given a constant $E > 1$ we just refine the mesh, $\mathcal{K}$, for the transport problem and return to 1. For example we can refine a certain percentage (refinement level) of the element according to the element indicators I_1^K.
10. If $I_2 > EI_1$ we just refine the mesh for the elliptic problem, $\mathcal{Q}$, according to I_2^K and return to 1.
11. If non of these hold we refine both $\mathcal{K}$ and $\mathcal{Q}$ and return to 1.

We do not consider adaptivity in time in this paper. However it will be very straight forward to do so following the work of Johnson et. al. [13, 12]. In the next section we will show a numerical example where we follow the steps of the adaptive algorithm closely.

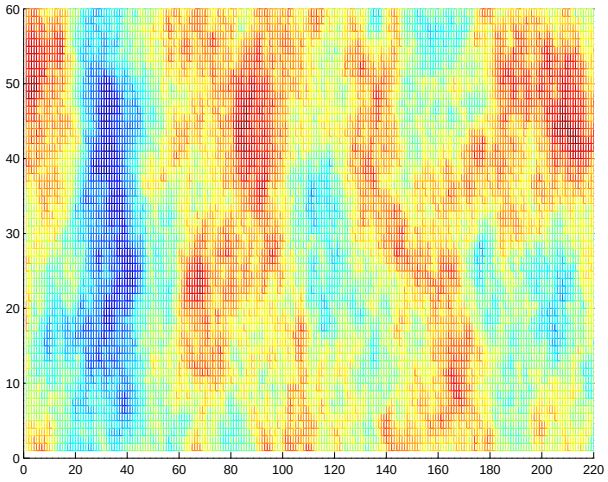


Figure 2: The permeability a used in the numerical example is taken from the Tenth SPE Comparative Solution Project. It is the top layer in the Tarbert formation in log-scale. $\max a / \min a = 1.53 \cdot 10^6$

5 Numerical Examples

We will study a simplified model of flow in petroleum reservoirs. The permeability a in the example is taken from the Tenth SPE Comparative Solution Project and is represented as a piecewise constant function on a 220×60 grid. We use a two dimensional slice of data representing the top layer of the Tarbert formation, see Figure 2. We start to solve the pressure equation (2.10) on a 220×60 block rectangular grid with $f = 1$ for $|(x, y)| \leq 0.025$ and $f = -1$ for $|(x, y) - (1, 1)| \leq 0.025$. In Figure 3 we see the flux $-\mathbf{\Sigma}$ to the left and the pressure U to the right. Using this vector field as a convection

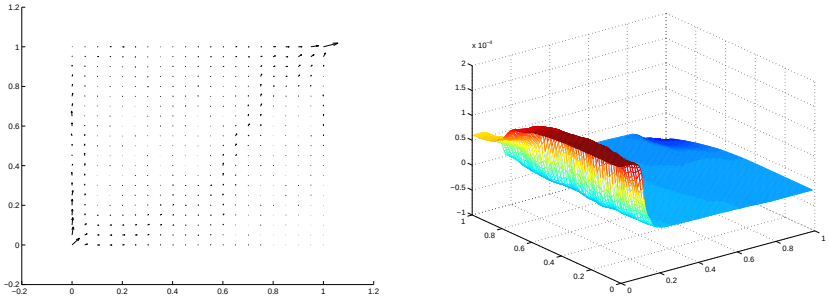


Figure 3: The solution to the pressure equation after one iteration.

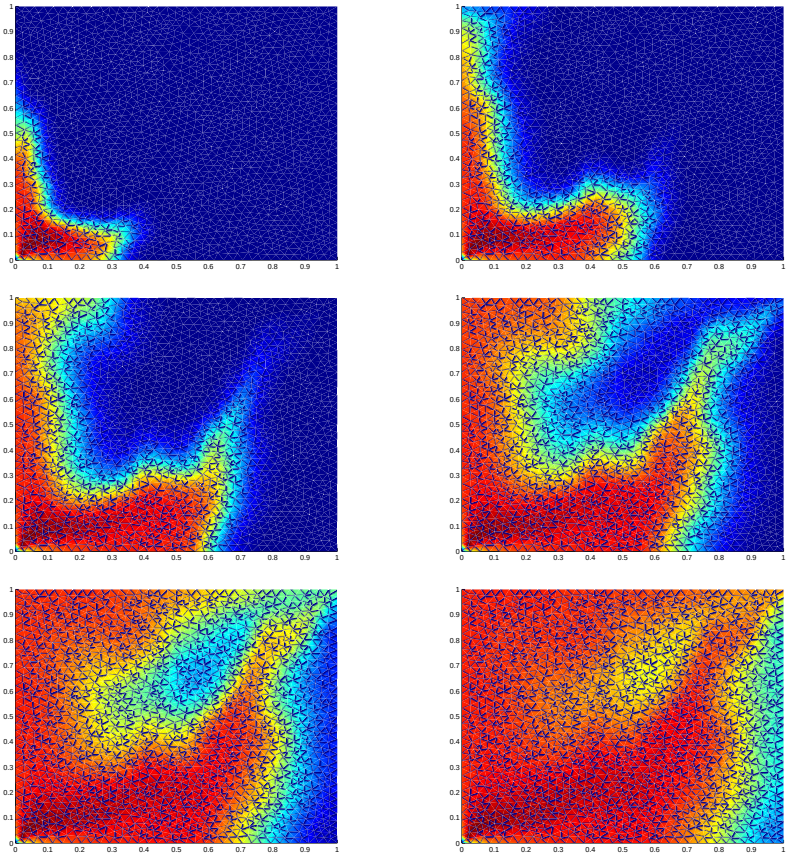


Figure 4: The solution to the transport problem after one iteration.

term we solve the transport problem, see equation (2.8), on a quasi uniform triangular mesh with mesh parameter $h \approx 0.03$ and $g = max(f, 0)$, see Figure 4 for solutions at six different times. We see how the water is injected in the lower left corner and how it follows the vector field from Figure 3 on its way to the producer in the upper right corner. We let $\epsilon = 3 \times 10^{-6}$ and use streamline diffusion to get stabilized solutions.

By using different kind of meshes for the two equations we want to show that the method of using coupled dual problems is very general. In many engineering applications each solver is its own black box and perhaps only some of the solvers supports element split or other local methods for improving the solution. In this example we will assume that the transport solver uses a triangular mesh that we can split to improve the solution locally while the elliptic solver does not support local mesh refinement. However

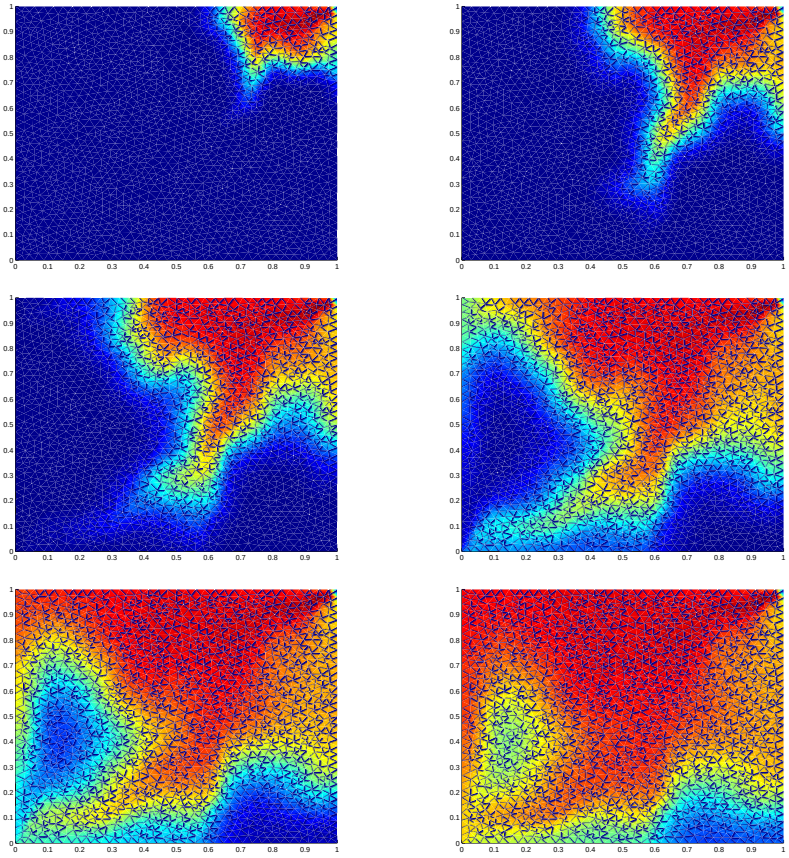


Figure 5: The dual solution to the transport problem after one iteration.

we can of course decrease the mesh size globally also for the elliptic solver.

Following the steps of the adaptive algorithm in Section 4 we are now ready to calculate the dual solutions. First we need to choose a right hand side ψ. In the oil reservoir application we want to measure the water concentration at the producer located in the upper right corner of the domain. We choose ψ to be the Dirac delta measure in the upper right corner, $x = y = 1$, $\psi = \delta_{(1,1)}$. This means that we control the quantity,

$$\int_0^T (e, \psi)\, dx = \int_0^T e(1, 1, t)\, dx. \tag{5.1}$$

In Figure 5 we see the dual transport solution at six different times. The dual problem indicates which areas that contributes most to the error. We plot the solution backward in time and we can see how it propagates from

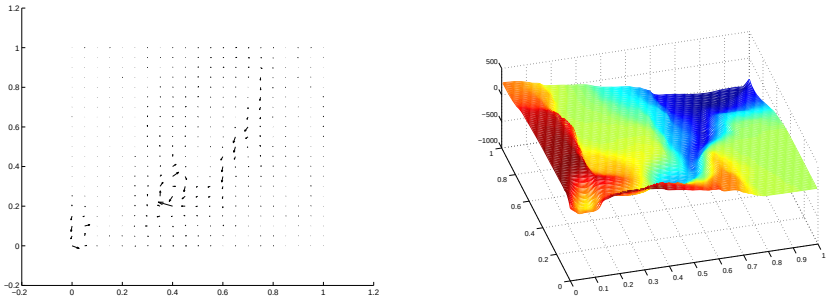


Figure 6: The dual solution to the pressure equation after one iteration.

the producer and through the domain. The dual problem is solved using the same method as the primal on the same mesh. Given an approximation to the primal and the dual we are ready to calculate the right hand side for the dual elliptic problem. We remember that,

$$\boldsymbol{\omega} = \int_0^T c\nabla\pi\phi\, dx \approx \int_0^T C\nabla\Phi\, dx. \tag{5.2}$$

The dual solution we get using this right hand side can be seen in Figure 6. To the left we see the flux. Remember that $\nabla \cdot \boldsymbol{\chi} = 0$ so we have no production in this variable, just a curl part this gives characteristic eddies. One can see that we have strong eddies at the producer and along the main flow path lower of the diagonal of the domain. Especially we see large magnitudes of the solution where the primal to the transport problem C changes direction. To the right we see the gradient part of the solution η. We can clearly see how it follows the flow from the solution $\boldsymbol{\Sigma}$ and therefore also C and Φ.

We are now ready to start calculating the error indicators at step 6 in the adaptive algorithm. We start by I_1^K and divide it into two parts. The first part is the residual denoted $\rho_K(C)$ throughout the paper, see Figure 7. The residual gets it main contributions along the front of the wave C. The second part is the weights from the dual solution, $\triangle t\|\dot{\phi}\|_K + h\|\nabla\bar{\phi}\|_K$, see Figure 8. Again we see that also the weight part of the error indicator follows the front, this time of the dual solution Φ.

We also have an error indicator associated with the elliptic problem. In Figure 9 to the left we see the residual part of the primal solution $\|\nabla U^* - \boldsymbol{\Sigma}/a\|_K + h^{-1/2}\|[U^*]\|_{\partial K}$ in log-scale. Remembering the permeability a in Figure 2 we recognize the difficult vertical area to the left of the domain where we also get large error indicators. We also not large values very close to the injector and producer. To the right in Figure 9 we have the other

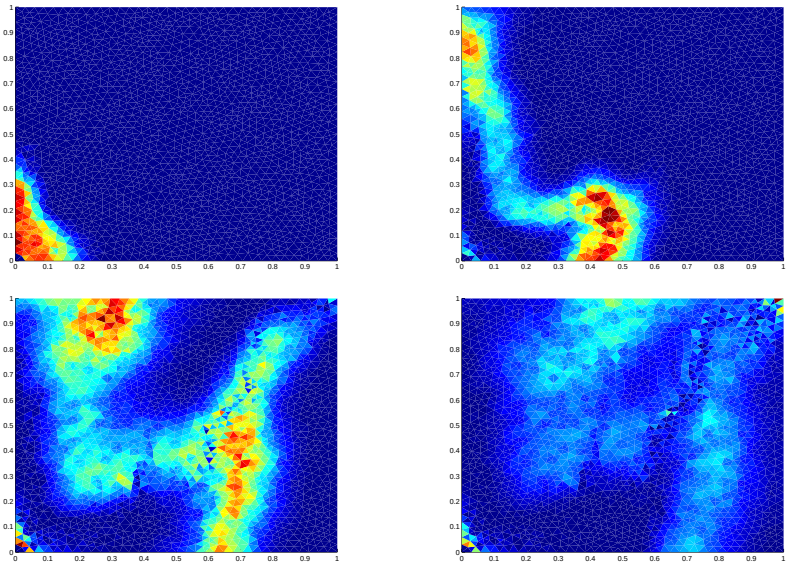


Figure 7: The residual ρ after the first iteration.

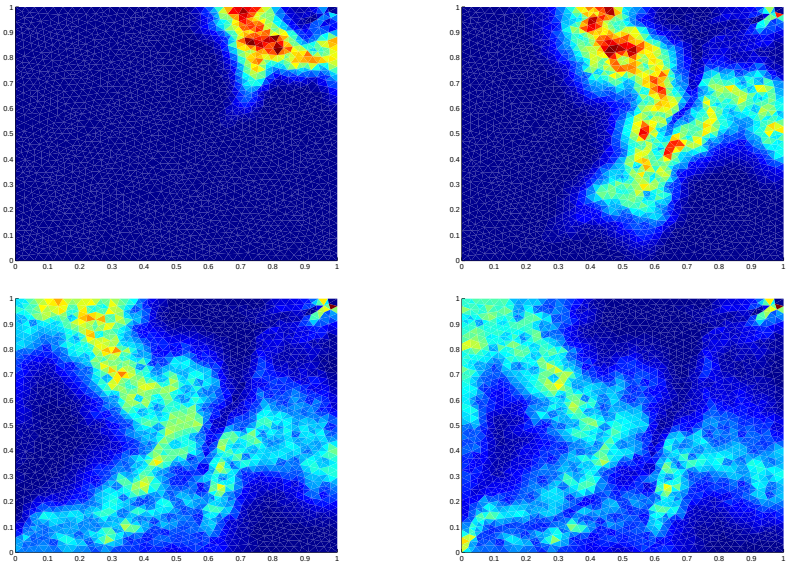


Figure 8: The weight function $\triangle t\|\dot{\phi}\|_K + h\|\nabla\phi\|_K$ backwards in time.

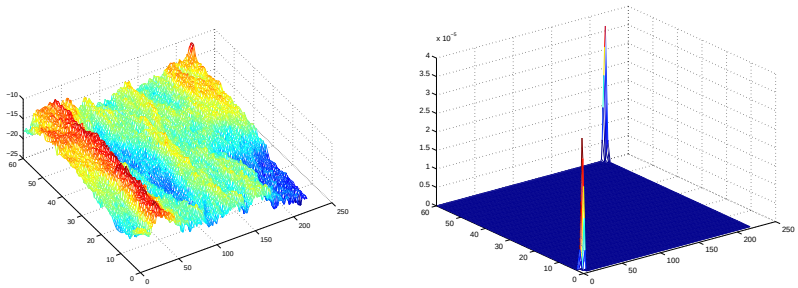


Figure 9: The two residual parts of the elliptic equation. To the left $\|\nabla U^* - \mathbf{\Sigma}/a\|_K + h^{-1/2}\|[U^*]\|_{\partial K}$ in log-scale and to the right $h\|f + \nabla \cdot \mathbf{\Sigma}\|$.

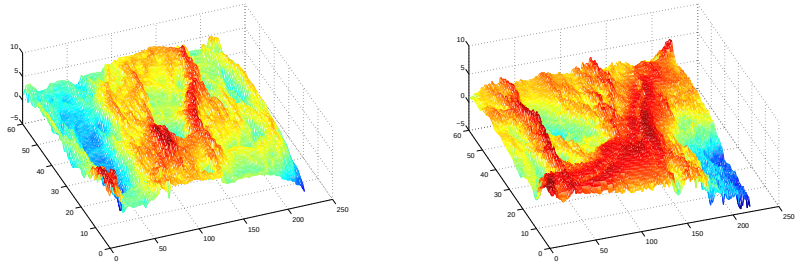


Figure 10: The weight functions for the elliptic equation in log-scale. To the left $\|\boldsymbol{\chi}\|_K$ and to the right $\|\nabla\eta\|_K$.

contribution to the residual, $h\|f + \nabla \cdot \mathbf{\Sigma}\|$. Since f is zero almost everywhere and the method is locally conservative we mostly get zero contribution from this part. However we also here get contributions close to the injector and producer where f is non-zero.

In Figure 10 we see the functions $\boldsymbol{\chi}$ to the left and $\nabla\eta$ to the right. Here it is obvious how the influence of the transport equation affects the error indicator of the elliptic problem. If we just would have used the residual from Figure 9 as an indicator for refinement we would lose all information that is associated with which output quantity we seek to minimize. This information is stored in the weights and needs to be taken into account.

We are now ready to use this error indicators to refine the mesh and iterate until we are satisfied with the solution. We use a refinement level of 15% for the transport equation and 100% for the elliptic equation, since we can not refine this mesh locally. We let $E = 2.5$ in step 8 and 9 of the adaptive algorithm which means that if the transport error is larger then 2.5 times the elliptic error, we only refine the mesh for the transport problem

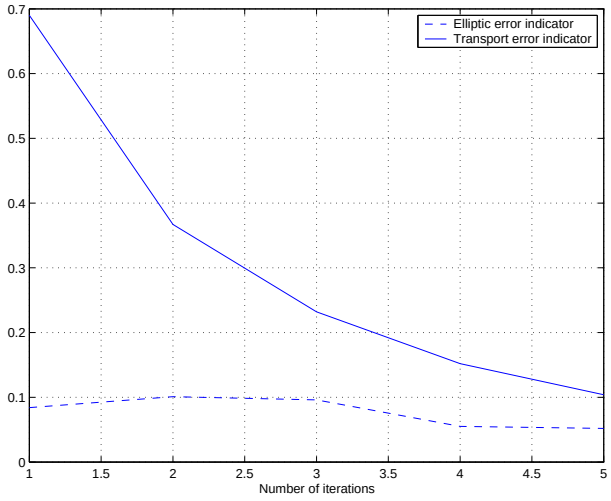


Figure 11: The error indicators for the transport and the elliptic problem at each iteration. We see a steady decay for the transport part as we expect since the mesh is refined at each iteration. For the elliptic indicator we see a drop after three iterations. This is because we refine the mesh uniformly after this iteration. This is the only modification the algorithm does to the mesh of the elliptic problem.

and vice versa.

We perform five iterations of the adaptive algorithm. In Figure 11 we plot the error indicators I_1 and I_2 after each iteration. We see that the largest contribution to the error comes from the transport problem after the first three iterations. We refine the triangular mesh by 15% at each iteration and the indicator decreases. Since we do not refine the mesh for the elliptic problem this indicator is fairly constant. At the fourth iteration we no longer have $I_1 > 2.5I_2$ so we refine both meshes. For the elliptic solver this means a uniform refinement. As seen from the figure this decreases the error indicator both for the elliptic and transport problem. At the last iteration we again just refines the triangular mesh. In Figure 12 we plot the five triangular meshes together with the initial rectangular mesh. At the fourth iteration the rectangular mesh will be uniformly refined but it will of course have the same structure so this mesh is not included in the picture. We see clearly how the mesh follows the two main paths of the flow.

The final solution on the last mesh can be seen in Figure 13. Analysis of the error indicators is one way of evaluate the method. An other way is to see how the solutions seems to converge as we put more and more

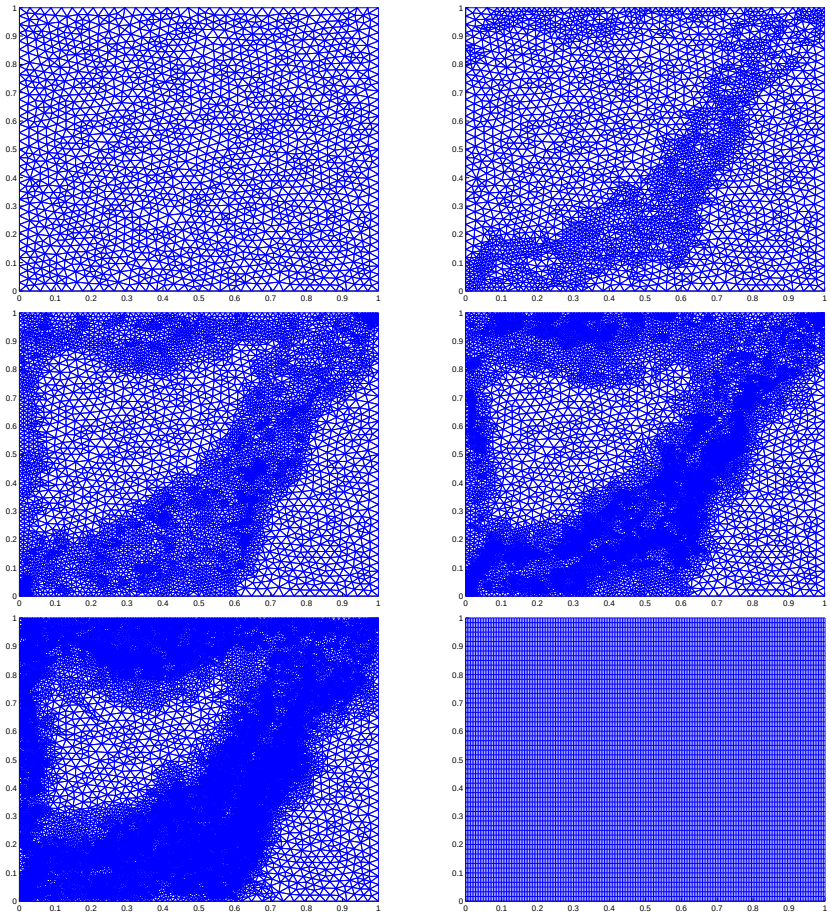


Figure 12: Five meshes for the transport problem after each of the five iterations. In the lower right corner the first mesh for the elliptic problem. A uniform refinement of the mesh for the elliptic problem is performed after the third iteration but this uniformly refined rectangular mesh is not present in the figure.

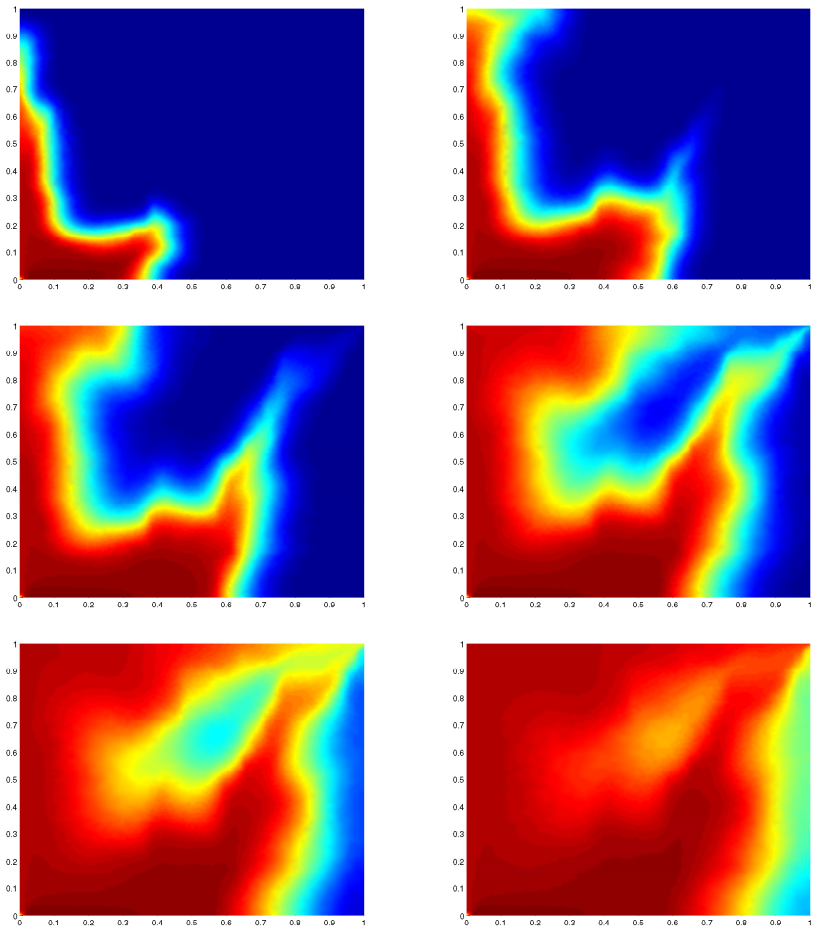


Figure 13: The final solution to the transport problem after five iterations.

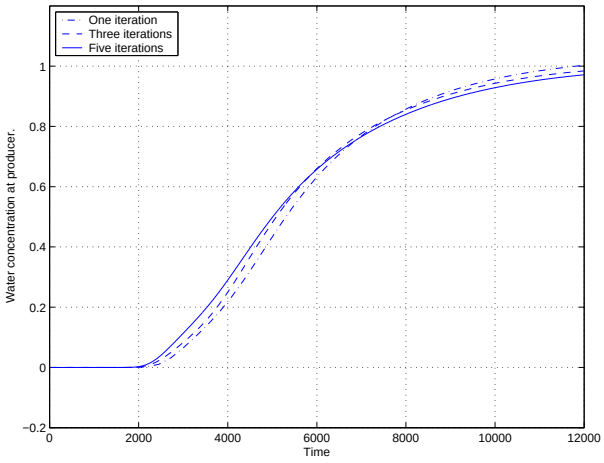


Figure 14: The water concentration at the producer at different times for approximations after one, three, and five iterations.

computational effort into solving the problem. Since we are interested in a specific output $\int_0^T C(1,1,t)\,dx$ we will now look how this quantity varies in the different approximations.

In Figure 14 we plot $C(1,1,t)$ after one three and five iterations. As we expect we have no water (pure oil) until the front reaches the producer and then the concentration of water increases to finally reach one. We zoom in on two parts of this figure. First we study the time when the front reaches the producer, see Figure 15 (left). We see that after more iterations we get steeper and steeper solutions. The fourth solution seems to be very close to the fifth. The second zoom is a bit later in the calculation. Again the fourth

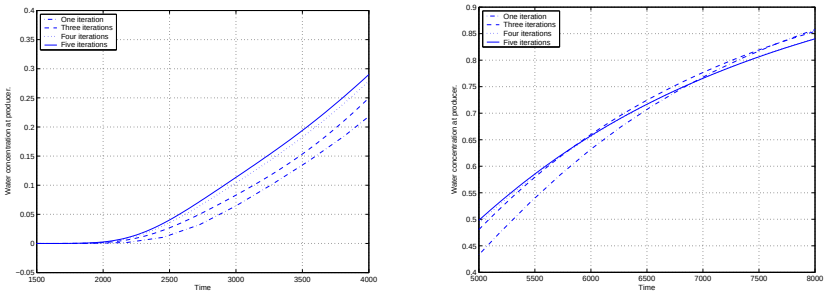


Figure 15: The water concentration at the producer zoomed around two specific times for four approximations.

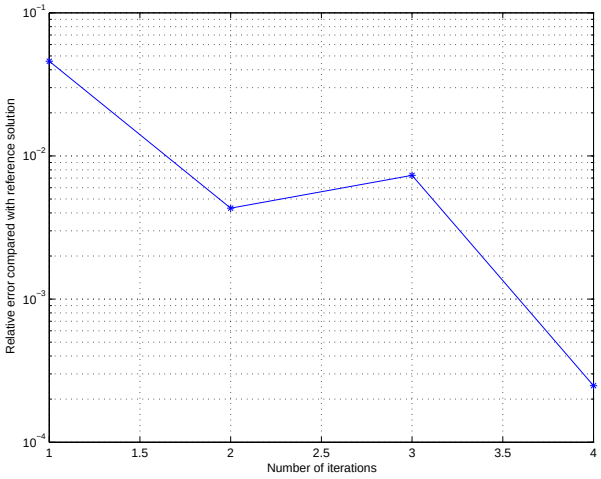


Figure 16: We choose the fifth approximation as reference to compare the other four approximations with. We calculate the relative error in concentration of water at the producer integrated over time.

and fifth iterates are very close.

We now use the fifth iterate as a reference solution and study the how close the prior iterates are to the last one. In Figure 16 we plot,

$$|\int_0^T C^i(1,1,t)\,dx - \int_0^T C^5(1,1,t)\,dx|, \tag{5.3}$$

for $i = 1, 2, 3, 4$, i indicating the four prior iterates. We see a big drop in the error between the first and second iteration then a small increase followed by an other big drop. Even though the decrease in error is not monotone we see orders of magnitude lower error for the solutions using refine meshes compared to the quasi regular meshes in the first iteration. Remember that just 15% of the elements where refined at each iteration.

References

[1] J.E. Aarnes, T. Gimse, and K.-A. Lie, *An introduction to the numerics of flow in porous media using Matlab,* Geometrical Modelling Numerical Simulation and Optimization, Industrial Mathematics at SINTEF. Springer Verlag, 2005.

[2] J.E. Aarnes, K.-A. Lie, and A. B. Rustad, *Modelling of Multiscale Structures in Flow Simulations for Petroleum Reservoirs,* submitted.

[3] R. A. Adams, *Sobolev spaces,* volume 65 of Pure and Applied Mathematics, Academic Press, New York, 1975.

[4] T. Arbogast and S. L. Bryant, *A two-scale numerical subgrid technique for waterflood simulations,* SPE J., 2002.

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

[6] W. Bangerth and R. Rannacher *Adaptive Finite Element Methods for Differential Equations* Birkhäuser, 2003.

[7] R. Becker, P. Hansbo, and M. G. Larson, *Energy norm a posteriori error estimation for discontinuous Galerkin methods,* Comput. Methods Appl. Mech. Engrg., 192, 723-733, 2003.

[8] J. H. Bramble and J. Xu, *A local post-processing technique for improving the accuracy in mixed finite-element approximations,* SIAM J. Numer. Anal., 26, 1267-1275, 1989.

[9] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods,* Springer Verlag, 1994.

[10] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods,* Springer Verlag, 1991.

[11] C. Dawson, S. Sun, and M. F. Wheeler, *Compatible algorithms for coupled flow and transport,* Comput. Methods Appl. Mech. Engrg., 193, 2565-2580, 2004.

[12] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Introduction to adaptive methods for differential equations,* Acta Numerica, 1-54, 1995.

[13] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, *Computational differential equations,* Studentlitteratur, 1996.

[14] K. Eriksson and C. Johnson, *Adaptive streamline diffusion finite element methods for stationary convection-diffusion problems,* Math. Comp., 60, 167-188, 1993.

[15] C. Johnson, *Numerical solution of partial differential equations by the finite element method,* Studentlitteratur, 1987.

[16] S. Larsson and V. Thomée, *Partial Differential Equations with Numerical Methods,* Springer Verlag, 2005.

[17] C. Lovadina and R. Stenberg, *Energy norm a posteriori error estimates for mixed finite element methods,* Helsinki University of Technology. Research Report A473. September 2004.

[18] A. Målqvist and M. G. Larson, *A Posteriori Error Estimates in the Energy Norm for Mixed Finite Element Methods,* submitted.

[19] J.-C. Nédélec, *A new family of mixed finite elements in* $\mathbf{R}^3$*,* Numer. Math., 50, 57-81, 1986.

[20] P. Raviart and J. Thomas, *A mixed finite element method for second order elliptic problems,* in Mathematical Aspects of the Finite Element Method. Lecture Notes in Math. 606, Springer-Verlag, 292-315, 1977.

[21] R. Stenberg, *Postprocessing schemes for some mixed finite elements,* RAIRO Modél. Math. Anal. Numér., 25, 151-167, 1991.