UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Convergence of goal-oriented adaptive finite element methods

Permalink

https://escholarship.org/uc/item/1cm230n5

Author

Pollock, Sara

Publication Date

2012

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Convergence of Goal-Oriented Adaptive Finite Element Methods

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Mathematics with a specialization in Computational Science

by

Sara Pollock

Committee in charge:

Professor Michael Holst, Chair
Professor Randolph Bank
Professor Jurijs Bazilevs
Professor David J. Benson
Professor Philip E. Gill

2012

Copyright
Sara Pollock, 2012
All rights reserved.

The dissertation of Sara Pollock is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2012

DEDICATION

In memory of my grandmother, Yemema P. Seligson, 1915-2009.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank Michael Holst, my advisor and mentor for bringing me here to UCSD, supporting me throughout my time here and introducing me to this beautiful field. I thank my committee members, most of whom I have had the chance to work with and all of whom have provided insight into this work. I thank my collaborator Yunrong Zhu for sharing his precise understanding of the material and being very enjoyable to work with.

I also thank my mentors from New Mexico, Vageli Coutsias and Sally Seidel for their friendship and enthusiasm as I've worked on this degree. Vageli has inspired me to be a mathematician since I met him as an undergraduate at UNM and he continues to do so. Finally, I thank my friends for their encouragement and support and for listening to me talk about math and weightlifting.

Chapter 2, in full, has been submitted for publication of the material as it may appear in Applied Journal of Numerical Mathematics, 2011, M. Holst and S. Pollock, Elsevier, 2011. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, has been submitted for publication of the material as it may appear in SIAM Journal of Numerical Analysis, 2012, M. Holst, S. Pollock, Y. Zhu, SIAM, 2012. The dissertation author was the primary investigator and author of this paper.

VITA

EDUCATION

2012 Ph. D. in Mathematics with a specialization in Computational Science, University of California, San Diego.

2008 M. S. in Applied Mathematics, University of Washington, Seattle.

2007 B. S. in Mathematics *summa cum laude*, University of New Mexico, Albuquerque.

1998 Studio Art Diploma, School of the Museum of Fine Arts, Boston, MA.

RESEARCH

2009-2012 *Goal-oriented adaptive finite element methods for linear and nonlinear problems*, advisor: Michael Holst, University of California, San Diego.

2006-2007 *Topological mapping of chemical space of small (drug-like) molecules*, and *Geometric mapping of conformation space of cyclooctane*, supervisors: Evangelos Coutsias and Tudor Oprea, University of New Mexico, Albuquerque.

2004-2005 *Data acquisition software for laser-scanning microscope*, supervisor: Wolfgang Rudolph, University of New Mexico, Albuquerque.

TEACHING

2011 Associate Instructor, University of California, San Diego.

2008-2010 Graduate Teaching Assistant, University of California, San Diego.

2007 Graduate Teaching Assistant, University of Washington, Seattle.

2004-2005 Undergraduate Teaching Assistant, University of New Mexico.

PUBLICATIONS

M. Holst, S. Pollock, and Y. Zhu, "Convergence of goal-oriented adaptive finite element methods for semilinear problems", *submitted for publication*, 2012. arXiv:math.NA/1203.1381.

M. Holst and S. Pollock, "Convergence of goal-oriented adaptive finite element methods for nonsymmetric problems", *submitted for publication*, 2011. arXiv:math.NA/1108.3660.

W.M. Brown, S. Martin, S.N. Pollock, E.A. Coutsias, J.P. Watson, "Algorithmic dimensionality reduction for molecular structure analysis", *J. Chem. Phys.*, 129(6), 064118 (pp.1-13) (2008), (web release, 8/14/2008).

S.N. Pollock, E.A. Coutsias, M.J. Wester and T.I. Oprea, "Scaffold topologies I: exhaustive enumeration up to eight rings", *J. Chemical Information and Modeling*, 2008 (Jul. 8), 48(7), 1304-1310, (web release 7/8/2008).

M.J. Wester, S.N. Pollock, E.A. Coutsias, T.K. Allu, S. Muresan and T.I. Oprea, "Scaffold topologies II: analysis of chemical databases", *J. Chemical Information and Modeling*, 2008 (Jul. 8), 48(7), 1311-1324, (web release, 7/8/2008).

ABSTRACT OF THE DISSERTATION

Convergence of Goal-Oriented Adaptive Finite Element Methods

by

Sara Pollock

Doctor of Philosophy in Mathematics with a specialization in Computational Science

University of California, San Diego, 2012

Professor Michael Holst, Chair

In this thesis we discuss convergence theory for goal-oriented adaptive finite element methods for second order elliptic problems. We develop results for both linear nonsymmetric and semilinear problems. We start with a brief description of the finite element method applied to these problems and some basic error estimates. We then provide a detailed error analysis of the method as described for each problem. In each case, we establish convergence in the sense of the quantity of interest with a goal-oriented variation of the standard adaptive finite element method using residual-based indicators.

In the linear case we establish the adjoint as the appropriate differential operator for the dual problem. We establish contraction of the quasi-error for each of the primal and dual problems yielding convergence in the quantity of interest. We follow these results with a complexity analysis of the method. In the semilinear case we introduce

three types of linearized dual problems used to establish our results. We give a brief summary of a priori estimates for this class of problems. After establishing contraction results for the primal problem, we then provide additional estimates to show contraction of the combined primal and dual system, yielding convergence of the goal function. We support these results with some numerical experiments.

Finally, we include an appendix outlining some common methods used in *a posteriori* error estimation and briefly describe iterative methods for solving nonlinear problems.

Chapter 1

Introduction

1.1 Background and overview of research

We start with an overview of recent results in adaptive and goal-oriented finite element methods. We then introduce the two main problems investigated in this thesis, and survey some basic tools used later in the analysis. In §1.2 we summarize the main results of Chapters 2 and 3.

1.1.1 Adaptive and goal-oriented methods

Adaptive finite element methods (AFEM) are those in which only select elements are refined at each iteration of the algorithm in an attempt to produce a more efficient overall approximation algorithm. In contrast, uniform methods globally refine the mesh at every step. Adaptive methods are effective at reducing the overall complexity or degrees of freedom in the problem and are of particular interest in problems with localized singularities. In this work, we are interested in extending convergence theory for AFEM to broader classes of error indicators, problems and adaptive algorithms.

In recent results, a number of different quantities have been shown to contract in adaptive settings. In [18], the total error, a linear combination of the energy error and an oscillation term, is shown to contract for the nonsymmetric elliptic problem. In [6] for symmetric elliptic problems and [16] for semilinear problems, contraction is established in terms of the quasi-error, a linear combination of energy error and error estimator. Additionally, in [6] decay of the total error is shown to achieve the optimal rate in terms of the number of degrees of freedom and the best approximation. More recently in [2] contraction is shown for semilinear problems using inexact solvers. In this last case, the form of error that contracts is similar to the quasi-error, except the error estimator is a function of the approximate rather than exact solution.

Goal-oriented methods are those designed to approximate a function of the solution $g(u)$ rather than the u, the weak solution to the PDE. The function $g(\cdot)$ is referred to as the goal function, and $g(u)$ the quantity of interest. The goal function may represent a physical quantity or characteristic of particular interest. The canonical example is an average over a subdomain or a line integral about its boundary. Goal-oriented methods are used in a number of applications [13] including pointwise *a posteriori* error estimation [20]. Goal-oriented adaptive finite element methods (GOAFEM) use error estimators based on a dual problem which involves the function g to guide the refinement towards an accurate approximation of the quantity of interest. Our results for convergence of GOAFEM are preceded by [19] for the scaled Laplacian. While we follow the basic goal-oriented framework outlined in that paper, the convergence proof for the linear nonsymmetric problem follows that in [6] and [16], establishing contraction of the quasi-error for both primal and dual problems. For the semilinear problem, our analysis departs from this framework substantially in order to establish contraction of a combined primal-dual quasi-error, where in this case the dual problem is coupled to the primal problem. The strong contraction results presented here are the first for goal-oriented methods applied to nonsymmetric elliptic and nonlinear problems.

The standard adaptive algorithm iterates the loop

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE} \,. \tag{1.1.1}$$

For goal-oriented adaptive methods, each iteration of the algorithm involves solving both a primal and a dual problem, calculating an estimate of the error on each element, marking an appropriate set of elements for refinement and refining the mesh for the next iteration. We employ a standard strong-form residual-based error indicator as in [6], [19] and [18] to estimate the error on each element at each iteration of the algorithm. Non-residual based error indicators may also be employed in adaptive methods, as in [7] for linear elliptic problems. For goal-oriented adaptive methods, much of the literature focuses on weak-form residual-based estimators, for example [20, 11, 10, 9, 22, 14, 5]. The advantage of the strong-form indicators in this context is their role in analytically determining the monotonic decrease of the (combined) quasi-error.

1.1.2 Problems considered

The goal of this research is to prove the convergence of goal-oriented adaptive finite element methods for a sequence of increasingly general problems. Here we present our first two steps in this process: A linear nonsymmetric problem and a semilinear problem. The next step and the focus of our current work is a coupled system of semilinear equations. We take as our starting point the recent results of Mario Mömmer and Rob Stevenson [19] for the symmetric diffusion problem. Our first results are for the elliptic problem given in strong form by

$$\begin{aligned} -\mathscr{L}(u) := -\nabla \cdot (A\nabla u) + b \cdot \nabla u + cu = f, \quad & \text{in } \Omega, \\ u = 0, \quad & \text{on } \partial\Omega, \end{aligned} \tag{1.1.2}$$

with weak formulation: find $u \in H_0^1(\Omega)$ such that

$$a(u,v) := \int_\Omega A\nabla u \cdot \nabla v + b \cdot \nabla uv + cuv\, dx = f(v), \quad \forall v \in H_0^1(\Omega) \tag{1.1.3}$$

where we follow the convention of [8] and associate the function $f \in L_2(\Omega)$ with its Riesz-representer

$$f(v) = \int_\Omega fv\, dx. \tag{1.1.4}$$

We make the following assumptions on the problem data:

Assumption 1.1.1. *Assumptions on nonsymmetric problem.*

1) $A : \overline{\Omega} \to \mathbb{R}^{d \times d}$*, Lipschitz and a.e. symmetric positive-definite.*
2) $b : \overline{\Omega} \to \mathbb{R}^d$*, with* $b_k \in L_\infty(\Omega)$ *, and b divergence-free.*
3) $c : \overline{\Omega} \to \mathbb{R}$*, with* $c \in L_\infty(\Omega)$*, and* $c(x) \geq 0$ *for all* $x \in \Omega$*.*
4) $f, g \in L_2(\Omega)$*.*

Following the analyses of [19], [16] and [6] we provide strong contraction results and complexity estimates for the goal-oriented problem of finding $g(u)$ where u is the solution of (1.1.3).

Next, we consider the goal-oriented problem for the semilinear PDE given in strong form by

$$\begin{aligned} -\mathscr{N}(u) := -\nabla \cdot (A\nabla u) + b(u) = f, \quad & \text{in } \Omega, \\ u = 0, \quad & \text{on } \partial\Omega, \end{aligned} \tag{1.1.5}$$

with weak formulation: find $u \in H_0^1(\Omega)$ such that

$$a(u,v) + \langle b(u), v\rangle = f(v), \quad \forall v \in H_0^1(\Omega) \tag{1.1.6}$$

where

$$a(u,v) := \int_\Omega A\nabla u \cdot \nabla v \tag{1.1.7}$$

and $\langle \cdot, \cdot \rangle$ denotes the L_2 inner-product. We make the assumptions

Assumption 1.1.2. *Assumptions on semilinear problem.*

1) $A : \Omega \to \mathbb{R}^{d\times d}$, Lipschitz and a.e. symmetric positive-definite.

2) $b : \Omega \times \mathbb{R} \to \mathbb{R}$ is monotone (increasing):

$$b'(\xi) \geq 0, \textit{ for all } \xi \in \mathbb{R}.$$

Here and in the remainder of the paper, we write $b(u)$ instead of $b(x,u)$ for simplicity.

3) $f, g \in L_2(\Omega)$.

4) There are $u_-, u_+ \in L_\infty$ which satisfy

$$u_-(x) < u(x), u_k(x) \leq u_+(x) \textit{ for almost every } x \in \Omega. \tag{1.1.8}$$

The iterate u_k in (1.1.8) is the solution to the discrete problem (1.1.22). The additional necessary property that b' is Lipschitz on $[u_-, u_+]$ is a consequence of (1.1.8). The *a priori* assumption (1.1.8) effectively places restrictions on the nonlinearity $b(\cdot)$ or possibly on the angles of the mesh with weaker restrictions on b as discussed in Chapter 3. In that chapter, we provide strong contraction results for a goal-oriented method for this problem. We also introduce an appropriate form of the error to establish these results.

1.1.3 Norms and Sobolev spaces

In the weak formulations (1.1.3) and (1.1.6) we seek solutions u and consider test functions v in the Sobolev space $H_0^1(\Omega)$. In this section we define this function space, its norm, and the relation to the energy norm for each problem.

The $W_p^k(\Omega)$ Sobolev norm is given by

$$\|u\|_{W_p^k(\Omega)}^p = \sum_{|\alpha|<k} \int_\Omega |D^\alpha u|^p \, dx, \quad 1 \le p < \infty, \tag{1.1.9}$$

$$\|u\|_{W_\infty^k(\Omega)} = \sum_{|\alpha|<k} \text{ess sup}_{\,\Omega} |D^\alpha u|, \tag{1.1.10}$$

where $D^\alpha u$ denotes the weak partial derivative of u with multi-index α [12]. Then the space $W_p^k(\Omega)$ is a subspace of $L_p(\Omega)$ and is given by

$$W_p^k(\Omega) = \left\{ u \in L_p(\Omega) \;\middle|\; \|u\|_{W_p^k} < \infty \right\}. \tag{1.1.11}$$

Equivalently, we may define the Sobolev space $W_p^k(\Omega)$ as the closure of the space $C^k(\Omega)$ under the $W_p^k(\Omega)$ norm given by (1.1.9) for $1 \le p < \infty$ [4]. The Sobolev spaces are all Banach spaces [12]. For $p = 2$ we denote

$$H^k(\Omega) := W_2^k(\Omega) \tag{1.1.12}$$

where the H^k are Hilbert spaces. Finally, denote

$$H_0^1(\Omega) = \left\{ u \in H^1(\Omega) \;\middle|\; u\big|_{\partial\Omega} = 0 \right\} \tag{1.1.13}$$

where the restriction to the boundary is meant in the sense of the trace as in [12].

Along with the native H^1-norm, we also make use of the energy norm for each problem. Many of the *a posteriori* estimates are developed with respect to the energy norm. For each of (1.1.3) and (1.1.6) define

$$|||v|||^2 := a(v,v). \tag{1.1.14}$$

In each case this norm is seen to be induced by the symmetric part of the problem. In sections 2.2.1 (respectively 3.2), the energy norm for each case is seen to be equivalent to the native norm. In particular, there is a continuity constant $M_{\mathscr{E}}$ with

$$a(u,v) \leq M_{\mathscr{E}} \|u\|_{H^1} \|v\|_{H^1}, \text{ for all } u,v \in H_0^1(\Omega) \tag{1.1.15}$$

and a coercivity constant $m_{\mathscr{E}}$ with

$$a(v,v) \geq m_{\mathscr{E}}^2 \|v\|_{H^1}^2, \text{ for all } v \in H_0^1(\Omega), \tag{1.1.16}$$

yielding

$$m_{\mathscr{E}}^2 \|v\|_{H^1}^2 \leq a(v,v) \leq M_{\mathscr{E}} \|v\|_{H^1}^2. \tag{1.1.17}$$

1.1.4 Finite element methods

We consider approximating the solution u to (1.1.3) (respectively (1.1.6)) in the sequence of nested spaces

$$\mathbb{V}_0 \subseteq \mathbb{V}_1 \subseteq \mathbb{V}_2 \subseteq \ldots \subseteq \mathbb{V} = H_0^1(\Omega) \tag{1.1.18}$$

where each $\mathbb{V}_k$ is finite dimensional. A general discussion of finite element spaces and the finite element method may be found in a number of texts, including [17], [4] and [8]. For the results contained here, let $\mathscr{T}_0$ a conforming triangulation of the problem domain $\Omega \subset \mathbb{R}^d$, $d = 2$ or 3. We assume Ω itself is a polyhedral domain, and that $\mathscr{T}_0$ captures the boundary exactly. We consider refining the mesh by the method of newest vertex bisection [3], creating a sequence of shape-regular conforming meshes which preserve the smallest-angle condition. We define the finite element spaces by n-degree polynomials over each element, employing the notation

$$\mathbb{V}_T = H_0^1(\Omega) \cap \mathbb{P}_n(T), \quad \mathbb{V}_{\mathscr{T}} := H_0^1(\Omega) \cap \prod_{T \in \mathscr{T}} \mathbb{P}_n(T), \quad \text{and} \quad \mathbb{V}_k := \mathbb{V}_{\mathscr{T}_k}. \tag{1.1.19}$$

From the continuous linear problem (1.1.3), we approximate the solution u by defining the discrete problem: Find $u_k \in \mathbb{V}_k$ such that

$$a(u_k, v) = f(v), \quad \forall v \in \mathbb{V}_k. \tag{1.1.20}$$

The solution u_k to (1.1.20) is found by solving a linear system assembled from each element $T \in \mathscr{T}_k$.

The discrete space $\mathbb{V}_k$ is spanned by a finite-dimensional piecewise polynomial basis $\{\varphi_j\}_{j=1}^{M_k}$. Each φ_j a local basis function of $\mathbb{V}_T$, $T \in \mathscr{T}_k$ can be mapped to a global reference domain $\hat{T}$ with corresponding basis function $\hat{\varphi}_j$. As the basis functions for each local $\mathbb{V}_T$, $T \in \mathscr{T}_k$ map back to the same reference element, we ultimately consider only N basis functions $\{\hat{\varphi}_j\}_{j=1}^{N}$ corresponding to the basis functions with nonzero support over a given (interior) element. In the particular case of polynomials of degree n in dimension d we have $N = \binom{n+d}{d}$.

Writing u as an expansion in basis functions, consider $v = \varphi_i,\ i = 1, \ldots, M_k$. By linearity of $a(\,\cdot\,,\,\cdot\,)$

$$a\left(\sum_{j=1}^{M_k} \alpha_j \varphi_j, \varphi_i\right) = \sum_{j=1}^{M_k} \alpha_j a\left(\varphi_j, \varphi_i\right) = f(\varphi_i),\ i = 1, \ldots, M_k. \tag{1.1.21}$$

The basis functions φ_j are chosen with limited support so the matrix $a(\varphi_j, \varphi_i)$, $i, j = 1, \ldots, M_k$ is highly sparse, enabling the system produced by (1.1.21) to be solved efficiently by an iterative method [15], [1] (see, for instance [21]). In practice, this system is determined by assembling local element matrices over all $T \in \mathscr{T}_k$. The calculation of $a(\varphi_j, \varphi_i),\ i, j = 1, \ldots, N$, is performed on the reference domain $\hat{T}$ via $a(\hat{\varphi}_j, \hat{\varphi}_i)$. As the reference basis functions do not change throughout the adaptive algorithm (here, we enrich the approximation spaces by refining the elements, not through seeking solutions to higher-order polynomial approximations), this calculation needs only to be performed once.

For the case of a nonlinear problem as in (1.1.6) we define the discrete problem as above

$$a(u_j, v) + \langle b(u_j), v \rangle = f(v), \quad \forall v \in \mathbb{V}_j. \tag{1.1.22}$$

However, the first equality in (1.1.21) will not hold. To handle this situation, the method as described above is applied to the linearized problem. The nonlinear problem (1.1.22) is then solved by Newton iteration as described in Apppendix C.

1.1.5 Galerkin orthogonality and Céa's lemma

One of the key approximation theorems in the finite element method is Céa's Lemma [8], which shows that up to a constant, the Galerkin solution is the best approximation in the discrete approximation space

$$\|u-u_k\|_{H^1} \le C \inf_{v\in\mathbb{V}_k} \|u-v\|_{H^1}. \tag{1.1.23}$$

For the linear problem, let u the solution to (1.1.3) and u_k the solution to (1.1.20). Then

$$a(u-u_k,v) = f(v) - f(v) = 0, \text{ for all } v \in \mathbb{V}_k. \tag{1.1.24}$$

This property is referred to as Galerkin orthogonality. To establish Céa's lemma, use the relations (1.1.15) and (1.1.16) between $a(\cdot,\cdot)$ and the native norm along with (1.1.24)

$$\begin{aligned} |||u-u_k|||_{H^1}^2 &\le m_{\mathscr{E}}^{-2} a(u-u_k,u-u_k) \\ &= m_{\mathscr{E}}^{-2} a(u-u_k,u-v) \\ &\le \frac{M_{\mathscr{E}}}{m_{\mathscr{E}}^2} \|u-u_k\|_{H^1}\|u-v\|_{H^1}, \text{ for all } v\in\mathbb{V}_k. \end{aligned} \tag{1.1.25}$$

Canceling one factor of $\|u-u_k\|_{H^1}$ yields the result (1.1.23).

For the nonlinear problem, let u the solution to (1.1.6) and u_k the solution to (1.1.22). Galerkin orthogonality now takes the form

$$a(u-u_k,v) + \langle b(u)-b(u_k),v\rangle = f(v) - f(v) = 0, \text{ for all } v\in\mathbb{V}_k. \tag{1.1.26}$$

Here we establish (1.1.23) following the discussion in [23]. We start with two observations. First, by the monotonicity of b as in Assumption 1.1.2

$$\langle b(u)-b(v),u-v\rangle \geq 0 \text{ for all } u,v\in H_0^1(\Omega). \tag{1.1.27}$$

The second, there exists a constant L with

$$\langle b(u)-b(v),w\rangle \leq L\|u-v\|_{H^1}\|w\|_{H^1} \text{ for all } u,v\in H_0^1(\Omega)\cap[u_-,u_+] \text{ and } w\in H_0^1(\Omega). \tag{1.1.28}$$

The Lipschitz property of b in the native norm follows from the a priori assumption on u (1.1.8) and bounding the L_2 by the H^1 norm. Rearranging terms in (1.1.26) for the test-function $v-u_k\in\mathbb{V}_k$ and applying (1.1.27) and (1.1.28)

$$\begin{aligned} a(u-u_k,v-u_k) &= -\langle b(u)-b(u_k),v-u_k\rangle \\ &= \langle b(u_k)-b(u),v-u\rangle - \langle b(u_k)-b(u),u_k-u\rangle \\ &\leq \langle b(u_k)-b(u),v-u\rangle \\ &\leq L\|u-u_k\|_{H^1}\|u-v\|_{H^1}. \end{aligned} \tag{1.1.29}$$

Next, by coercivity (1.1.16) followed by (1.1.29) and continuity (1.1.15)

$$\begin{aligned} m_{\mathscr{E}}^2\|u-u_k\|_{H^1}^2 &\leq a(u-u_k,u-u_k) \\ &= a(u-u_k,u-v)+a(u-u_k,v-u_k) \\ &\leq a(u-u_k,u-v)+L\|u-u_k\|_{H^1}\|u-v\|_{H^1} \\ &\leq M\|u-u_k\|_{H^1}\|u-v\|_{H^1}+L\|u-u_k\|_{H^1}\|u-v\|_{H^1}. \end{aligned} \tag{1.1.30}$$

Canceling one factor of $\|u-u_k\|_{H^1}$

$$\|u-u_k\|_{H^1} \leq \frac{M+L}{m_{\mathscr{E}}^2}\|u-v\|_{H_1} \text{ for all } v\in\mathbb{V}_k, \tag{1.1.31}$$

establishing (1.1.23). Having established this basic approximation property of the finite element method, we now consider how to approximate $g(u)-g(u_k)$, the error in the goal function.

1.1.6 Duality methods

Next we discuss the formation of the dual problem and its relation to the residual. For goal-oriented adaptive methods, much of the literature focuses on weak-form residual-based estimators, for example [20, 11, 10, 9, 22, 14, 5]. Below, we show the relation between strong and weak forms of the residual and the different types of estimates they are used for.

For the goal-oriented problem we seek a functional $g(\cdot)$ of the weak solution u of the original or primal PDE. The dual problem is introduced to satisfy the relationship

$$g(e_k) = \langle R(u_k), z \rangle, \tag{1.1.32}$$

where $e_k = u - u_k$ is the error in the primal approximation and z is the solution to the dual problem. The residual $R(\cdot)$ is given for linear problem (1.1.2) by

$$R(v) = f + \mathscr{L}(v). \tag{1.1.33}$$

Integrating over the domain against a test-function yields a weak-form relation

$$\langle R(v), w \rangle = f(w) - a(v, w), \tag{1.1.34}$$

corresponding to the weak-form of the problem (1.1.3). Similarly, for the nonlinear problem (1.1.5) and (1.1.6)

$$R(v) = f + \mathscr{N}(v) \;\text{ and }\; \langle R(v), w \rangle = f(w) - \left(a(v, w) + \langle b(v), w \rangle\right). \tag{1.1.35}$$

In our analysis we use the convergence properties of both the primal and dual sequences of iterates, u_k and respectively z_k to bound the RHS of (1.1.32). In contrast, other methods as for example [11], [9], [22], and [5] approximate the dual solution z in a higher order space and project down to the primal finite element space to estimate the RHS of (1.1.32) by means of Galerkin orthogonality and the relation

$$\langle R(u_k), z \rangle = \langle R(u_k), z - \pi_k z \rangle \approx \langle R(u_k), \tilde{z} - \pi_k \tilde{z} \rangle, \tag{1.1.36}$$

where $\tilde{z}$ denotes an approximation of z and π_k is the projector onto $\mathbb{V}_k$. The RHS of (1.1.36) is a computable quantity and may be used as an error indicator either alone or in conjunction with a primal indicator independent of the dual problem or its solution. Techniques of this form have been numerically shown to effectively guide the refinement towards an accurate evaluation of the goal function [20, 11, 10, 9, 5]. These techniques have not, however, been proven to monotonically decrease the error in the goal function. On the other hand, following the methods in [19] for the symmetric linear problem, we can take (1.1.32) and bound the RHS by an energy-norm estimate of the form

$$|g(e_k)| \leq K \left(|||u - u_k|||^2 + |||z - z_k|||^2 \right). \tag{1.1.37}$$

Following the convergence framework in [16], and [6], we show the quantity on the RHS of (1.1.37) is bounded by a form of the error which is reduced at each iteration of the algorithm, proving convergence of the method.

For the linear (divergence-free) elliptic problem with symmetric diffusion coefficient as given by (1.1.3) the dual problem with $a^*(\cdot,\cdot)$ the formal adjoint of $a(\cdot,\cdot)$ given by the RHS of (1.1.38) satisfies the relationship (1.1.32). Integrating by parts on the convergence term gives

$$\begin{aligned} a(u,v) &:= \langle A\nabla u, \nabla v\rangle + \langle b \cdot \nabla u, v\rangle + \langle cu, v\rangle \\ &= \langle A\nabla v, \nabla u\rangle - \langle b \cdot \nabla v, u\rangle + \langle cv, u\rangle =: a^*(v,u). \end{aligned} \tag{1.1.38}$$

The dual problem is then defined as: Find $z \in H_0^1(\Omega)$ such that

$$a^*(z,v) = g(v), \text{ for all } v \in H_0^1(\Omega), \tag{1.1.39}$$

and by (1.1.38) obtain the property (1.1.32) by

$$g(e_k) = a^*(z,u) - a^*(z,u_k) = a(u,z) - a(u_k,z) = f(z) - a(u_k,z) = \langle R(u_k), z\rangle. \tag{1.1.40}$$

The convergence and complexity analysis for the goal-oriented method based on dual problem (1.1.39) is the topic of Chapter 2.

For nonlinear problems, we may not have a formal adjoint available. For the problem as given by (1.1.6) we form the dual by linearization. By the integral mean-value theorem [10] or equivalently generalized Taylor expansion [23]

$$b(u)-b(u_k)=\int_0^1 b'(u_k+t(u-u_k))dt(u-u_k)=\int_0^1 b'(tu+(1-t)u_k)dt(u-u_k) \tag{1.1.41}$$

yielding

$$b(u)-b(u_k)=\mathscr{B}_k(u-u_k), \quad \mathscr{B}_k:=\int_0^1 b'(tu+(1-t)u_k)dt. \tag{1.1.42}$$

Noting $\mathscr{B}_k=\mathscr{B}_k^*$ we define the dual problem: Find $z\in H_0^1(\Omega)$ such that

$$a(z,v)+\langle\mathscr{B}_k z,v\rangle=g(v), \text{ for all } v\in H_0^1(\Omega). \tag{1.1.43}$$

Then by (1.1.6), (1.1.42) and (1.1.43)

$$\begin{aligned} g(e_k)&=a(z,u-u_k)+\langle\mathscr{B}_k z,u-u_k\rangle \\ &=a(z,u-u_k)+\langle z,\mathscr{B}_k(u-u_k)\rangle \\ &=a(z,u-u_k)+\langle z,b(u)-b(u_k)\rangle \\ &=a(u,z)+\langle b(u),z\rangle-(a(u_k,z)+\langle b(u_k),z\rangle) \\ &=f(z)-(a(u_k,z)+\langle b(u_k),z\rangle)=\langle R(u_k),z\rangle, \end{aligned} \tag{1.1.44}$$

which again satisfies the relation (1.1.32). The dual problem as given by (1.1.42) and (1.1.43) suffers two major drawbacks with respect to determining convergence of an iterative method. First, the linearized operator $\mathscr{B}_k$ itself is not computable as it is a function of the exact solution u to (1.1.6). Second, it is also a function of a particular approximate solution, in this case the primal iterate u_k. The first problem may be dealt with by replacing the linearized dual operator as in (1.1.42) with the approximate dual operator $b'(u_k)$. The dual problem defined with this operator does not satisfy the relation (1.1.32); however, it does yield a computable dual problem for each discrete primal solution u_k. The second problem, an issue in the convergence analysis, may be handled by introducing a limiting dual operator $b'(u)$, which again is not computable and does not satisfy relation (1.1.32); however, it may be used in the analysis to define a suitable form of error to contract at each iteration. A convergence analysis which handles the error terms induced by using the approximate and limiting dual operators in place of the linearized dual operator for the semilinear problem is detailed in Chapter 3.

1.1.7 Estimators, quasi-error and contraction

In this section we introduce the appropriate form of error to show contraction at each iteration of the goal-oriented algorithm. We also outline the main steps in the standard contraction framework. We start with the error indicator based on the strong-form of the residual. The error indicator is given elementwise as

$$\eta_{\mathscr{T}}^2(v,T) := h_T^2 \|R(v)\|_{L_2(T)}^2 + h_T \|J_T(v)\|_{L_2(\partial T)}^2, \quad v \in \mathbb{V}_{\mathscr{T}}, \tag{1.1.45}$$

where the mesh diameter $h_T = |T|^{1/d}$ and the residual $R(\cdot)$ is given by (1.1.33) for the linear and (1.1.35) for the nonlinear primal problems. The dual residuals are defined analogously. The jump residual for primal and dual problems in all cases is

$$J_T(v) := [\![[A\nabla v] \cdot n]\!]_{\partial T}, \tag{1.1.46}$$

where *jump operator* $[\![\ \cdot \ ]\!]$ is given by

$$[\![\phi]\!]_{\partial T} := \lim_{t \to 0} \phi(x+tn) - \phi(x-tn), \tag{1.1.47}$$

and n is taken to be the appropriate outward normal defined piecewise on ∂T. The error estimator is given by the l_2 sum of indicators. For the Galerkin solution u_k we use the notation

$$\eta_k^2 = \sum_{T \in \mathscr{T}_k} \eta_T^2(u_k, T). \tag{1.1.48}$$

As in [16] and [6] for each of the primal and dual problems in the linear nonsymmetric case (1.1.3) and (1.1.39), we establish contraction of the quasi-error given by a scaled sum of energy error and error estimator

$$Q_k^2(u_k, \mathscr{T}_k) := |||u - u_k|||^2 + \gamma_p \eta_k^2 \ \text{ and } \ Q_k^2(z_k, \mathscr{T}_k) := |||z - z_k|||^2 + \gamma_d \zeta_k^2, \tag{1.1.49}$$

where $\gamma_p, \gamma_d > 0$ and ζ_k denotes the dual error estimator. For the semilinear case (1.1.6) and (1.1.43) we see contraction in the combined quasi-error

$$\bar{Q}_k(u_k, \hat{z}_k) := |||\hat{z} - \hat{z}_k|||^2 + \gamma \zeta_k^2(\hat{z}_k) + \pi |||u - u_k|||^2 + \pi \gamma_p \eta_k^2(u_k), \tag{1.1.50}$$

where γ, γ_p and $\pi > 0$. Here, the contraction is stated with respect to the limiting dual problem. As the dual problem is a function of the primal solution, the contraction of the dual is coupled to the contraction of the primal quasi-error.

In both linear and nonlinear cases, the contraction argument follows from combining three main estimates as in [16] and [6].

1) Quasi-orthogonality: There exists $\Lambda_G > 1$ such that

$$|||u - u_2|||^2 \leq \Lambda_G |||u - u_1|||^2 - |||u_2 - u_1|||^2.$$

2) Error estimator as upper bound on error: There exists $C_1 > 0$ such that

$$|||u - u_k|||^2 \leq C_1 \eta_k^2(u_k, \mathscr{T}_k), \quad k = 1, 2.$$

3) Estimator reduction: For $\mathscr{M}$ the marked set that takes refinement $\mathscr{T}_1 \to \mathscr{T}_2$, for positive constants $\lambda < 1$ and Λ_1 and any $\delta > 0$

$$\eta_2^2(v_2, \mathscr{T}_2) \leq (1+\delta)\{\eta_1^2(v_1, \mathscr{T}_1) - \lambda \eta_1^2(v_1, \mathscr{M})\} + (1+\delta^{-1})\Lambda_1 \eta_0^2 |||v_2 - v_1|||.$$

In both cases, these three estimates are shown for both primal and (limiting) dual problems. In the linear case, the primal and dual quasi-errors can be shown independently to contract. In the semilinear case, the contraction of the primal error is used as a fourth key estimate in the contraction of the combined quasi-error.

1.2 Summary of the papers

1.2.1 Paper 1

In the first paper (Chapter 2), we consider the linear PDE as given by (1.1.2). We show contraction of the quasi-error for the primal and dual problem, independently of one another. In particular, there is an $\alpha < 1$ with

$$|||u - u_{k+1}|||^2 + \gamma_p \eta_{k+1}^2 \leq \alpha^2 \left(|||u - u_k|||^2 + \gamma_p \eta_k^2 \right), \text{ and} \tag{1.2.1}$$

$$|||z - z_{k+1}|||^2 + \gamma_d \zeta_{k+1}^2 \leq \alpha^2 \left(|||z - z_k|||^2 + \gamma_d \zeta_k^2 \right). \tag{1.2.2}$$

Putting this together with the bound for the error in the goal function

$$|g(u) - g(u_k)| \leq 2|||u - u_k||| \, |||z - z_k||| \tag{1.2.3}$$

we establish convergence of the method. As in [16] and [18] we assume the standard initial mesh condition

$$h_0^s \|b\|_{L_\infty} C_* \mu_0^{-1/2} < 1, \tag{1.2.4}$$

for some $s \in (0,1]$ depending on the angles of $\partial\Omega$ where $\|b\|_{L_\infty}$ and μ_0 are constants derived from the problem data, C_* is a global constant and h_0 is the maximum mesh diameter in the initial mesh $\mathscr{T}_0$.

We include a brief discussion of approximation classes $\mathbb{A}_s$. Then assuming the primal solution $u \in \mathbb{A}_s$ and the dual solution $z \in \mathbb{A}_t$, we derive the quasi-optimal complexity estimate

$$\begin{aligned} \#\mathscr{T}_k - \#\mathscr{T}_0 \leq S(\theta) \Bigg\{ & M_p \left(1 + \frac{\gamma_p}{c_2}\right)^{1/2s} Q_k^{-1/s}(u_k, \mathscr{T}_k) \\ & + M_d \left(1 + \frac{\gamma_d}{c_2}\right)^{1/2t} Q_k^{-1/t}(z_k, \mathscr{T}_k) \Bigg\}. \end{aligned} \tag{1.2.5}$$

1.2.2 Paper 2

In the second paper (Chapter 3), we consider the semilinear problem as given by (1.1.5). We show the contraction of the combined quasi-error $\bar{Q}_k(u_k,\hat{z}_k)$

$$\bar{Q}^2_{k+1}(u_{k+1},\hat{z}_{k+1}) \leq \alpha_D^2 \bar{Q}^2_k(u_k,\hat{z}_k),\ \alpha_D < 1, \tag{1.2.6}$$

where $\bar{Q}_k(u_k,z_k)$ is given by (1.1.50). We derive a bound for the error in the goal function

$$|g(u)-g(u_j)| \leq \frac{1}{2}(1+K_1 h_0^{2s})|||u-u_j|||^2 + \frac{1}{2}(1+K_2 h_0^{2s}), \tag{1.2.7}$$

for global constants K_1 and K_2. Putting together (1.2.6) and (1.2.7) we establish convergence of the method. Here we make the initial mesh assumption

$$h_0^s B m_{\mathscr{E}}^{-1} C_* < 1, \tag{1.2.8}$$

for some $s \in (0,1]$ depending on the angles of $\partial\Omega$ where B and $m_{\mathscr{E}}$ are constants derived from the problem data, C_* is a global constant and h_0 is the initial mesh diameter as above. Lastly, we show some numerical experiments which support our theoretical results for the goal-oriented method.

References

[1] R. Bank. Course notes for math 272, University of California San Diego, 2010.

[2] R. Bank, M. Holst, R. Szypowski, and Y. Zhu. Convergence of AFEM for semilinear problems with inexact solvers, 2011.

[3] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.

[4] S. Brenner and L. Scott. *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, third edition, 2008.

[5] V. Carey, D. Estep, and S. Tavener. A posteriori analysis and adaptive error control for multiscale operator decomposition solution of elliptic systems I: Triangular systems. *SIAM J. Numer. Anal*, 47(1):740–761, 2009.

[6] J. M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.

[7] J. M. Cascon and R. H. Nochetto. Quasioptimal cardinality of AFEM driven by nonresidual estimators. *IMA Journal of Numerical Analysis*, 32(1):1–29, 2011.

[8] P. G. Ciarlet. *Finite Element Method for Elliptic Problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.

[9] D. Estep, M. Holst, and M. Larson. Generalized green's functions and the effective domain of influence. *SIAM J. Sci. Comput*, 26:1314–1339, 2002.

[10] D. Estep, M. Holst, and D. Mikulencak. Accounting for stability: A posteriori error estimates based on residuals and variational analysis. In *Communications in Numerical Methods in Engineering*, pages 200–2, 2001.

[11] D. Estep, M. G. Larson, and R. D. Williams. Estimating the error of numerical solutions of systems of reaction-diffusion equations. *Mem. Amer. Math. Soc.*, 146(696):101–109, 2000.

[12] L. C. Evans. *Partial Differential Equations (Graduate Studies in Mathematics, V. 19) GSM/19*. American Mathematical Society, 1998.

[13] M. Giles and E. Süli. Adjoint methods for pdes: *a posteriori* error analysis and postprocessing by duality. *Acta Numerica*, pages 145–236, 2002.

[14] T. Grätsch and K.-J. Bathe. Influence functions and goal-oriented error estimation for finite element analysis of shell structures. *International Journal for Numerical Methods in Engineering*, 63:709–736, 2005.

[15] M. Holst. Course notes for math 273b, University of California San Diego, 2010.

[16] M. Holst, G. Tsogtgerel, and Y. Zhu. Local and global convergence of adaptive methods for nonlinear partial differential equations, 2008.

[17] T. Hughes. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Prentice-Hall, Englewood Cliffs, NJ, 1987.

[18] K. Mekchay and R. Nochetto. Convergence of adaptive finite element methods for general second order linear elliptic PDE. *SINUM*, 43(5):1803–1827, 2005.

[19] M. S. Mommer and R. Stevenson. A goal-oriented adaptive finite element method with convergence rates. *SIAM J. Numer. Anal.*, 47(2):861–886, 2009.

[20] S. Prudhomme and J. T. Oden. On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors. *Computer Methods in Applied Mechanics and Engineering*, 176(1-4):313–331, 1999.

[21] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2nd edition, 2003.

[22] R. Söderlund and M. Larson. Adaptive finite element methods for two way coupled problems based on decoupled duals, 2011.

[23] I. Stakgold and M. J. Holst. *Green's Functions and Boundary Value Problems*. Wiley, Hoboken, NJ, 3rd edition, 2011.

Chapter 2

Convergence of Goal-Oriented Adaptive Finite Element Methods for Nonsymmetric Problems

Convergence of Goal-Oriented Adaptive Finite Element Methods for Nonsymmetric Problems

Michael Holst, and Sara Pollock

Abstract. In this article we develop convergence theory for a class of goal-oriented adaptive finite element algorithms for second order nonsymmetric linear elliptic equations. In particular, we establish contraction and quasi-optimality results for a method of this type for second order Dirichlet problems involving the elliptic operator $\mathscr{L}u = \nabla \cdot (A\nabla u) - b \cdot \nabla u - cu$, with A Lipschitz, almost-everywhere symmetric positive definite (SPD), with b divergence-free, and with $c \geq 0$. We first describe the problem class and review some standard facts concerning conforming finite element discretization and error-estimate-driven adaptive finite element methods (AFEM). We then describe a goal-oriented variation of standard AFEM (GOAFEM). Following the recent work of Mommer and Stevenson for symmetric problems, we establish contraction of GOAFEM. We also then show convergence in the sense of the goal function. Our analysis approach is signficantly different from that of Mommer and Stevenson, combining the recent contraction frameworks developed by Cascon et. al, by Nochetto, Siebert, and Veeser, and by Holst, Tsogtgerel, and Zhu. In the last part of the paper we perform a complexity analysis, and establish quasi-optimal cardinality of GOAFEM. We include an appendix discussion of the duality estimate as we use it here in an effort to make the paper more self-contained.

2.1 Introduction

In this article we develop convergence theory for a class of goal-oriented adaptive finite element methods for second order nonsymmetric linear elliptic equations. In particular, we report contraction and quasi-optimality results for a method of this type for the problem

$$-\nabla \cdot (A\nabla u) + b \cdot \nabla u + cu = f, \quad \text{in } \Omega, \tag{2.1.1}$$

$$u = 0, \quad \text{on } \partial\Omega, \tag{2.1.2}$$

with $\Omega \subset \mathbb{R}^d$ a polyhedral domain, $d = 2$ or 3, with A Lipschitz, almost-everywhere (a.e.) symmetric positive definite (SPD), with b divergence-free, and with $c \geq 0$. The standard weak formulation of this problem reads: Find $u \in H_0^1(\Omega)$ such that

$$a(u,v) = f(v), \quad \forall v \in H_0^1(\Omega), \tag{2.1.3}$$

where

$$a(u,v) = \int_\Omega A\nabla u \cdot \nabla v + b \cdot \nabla uv + cuv \, dx, \qquad f(v) = \int_\Omega fv \, dx. \tag{2.1.4}$$

Our approach is to first describe the problem class in some detail, and review some standard facts concerning conforming finite element discretization and error-estimate-driven adaptive finite element methods (AFEM). We will then describe a goal-oriented variation of standard AFEM (GOAFEM). Following the recent work of Mommer and Stevenson [10] for symmetric problems, we establish contraction of GOAFEM. We also show convergence in the sense of the goal function. Our analysis approach is signficantly different from that of Mommer and Stevenson [10], combining the recent contraction frameworks of Cascon et. al [4], of Nochetto, Siebert, and Veeser [11], and of Holst, Tsogtgerel, and Zhu [8]. We also give a complexity analysis, and establish quasi-optimal cardinality of GOAFEM.

The goal-oriented problem concerns achieving a target quality in a given linear functional $g\colon H_0^1(\Omega) \to \mathbb{R}$ of the weak solution $u \in H_0^1(\Omega)$ of the problem (2.1.3). For example, $g(u) = \int_\Omega \frac{1}{|\omega|}\chi_\omega u$, the average value of u over some domain $\omega \subset \Omega$. By writing down the adjoint operator, $a^*(z,v) = a(v,z)$, we consider the *adjoint* or *dual* problem: find $z \in H_0^1(\Omega)$ such that $a^*(z,v) = g(v)$, for all $v \in H_0^1(\Omega)$. It has been shown for the symmetric form ($b = 0$) of problem (2.1.1)–(2.1.2) with piecewise constant SPD diffusion cofficient A (and with $c = 0$), that by solving the *primal* and *dual* problems simultaneously, one may converge to an approximation of $g(u)$ faster than by approximating u then $g(u)$, when forcing contraction in only the primal problem [10]. We will follow the same general approach in order to establish similar goal-oriented AFEM results for nonsymmetric problems. However, in order to handle nonsymmetry, we will follow the technical approach in [9, 4, 8], and rely largely on establishing quasi-orthogonality. In particular, contraction results are established in [9, 4] for (2.1.1)–(2.1.2) in the case that A is SPD, Lipschitz or piecewise Lipschitz, b is divergence-free, and $c \geq 0$. In [8], quasi-orthogonality is used as the basis for establishing contraction of AFEM for two classes of nonlinear problems. As in these earlier efforts, relying on quasi-orthogonality will require that we assume that the initial mesh is sufficiently fine, and that the solution to the dual problem $a^*(w,v) = g(v),\ g \in L_2(\Omega)$ is sufficiently smooth, e.g. in $H^2_{\text{loc}}(\Omega)$.

Following [8], the contraction argument developed in this paper will follow from first establishing three preliminary results for two successive AFEM approximations u_1 and u_2, and then applying the Dörfler marking strategy:

1) Quasi-orthogonality (§2.3.1): There exists $\Lambda > 1$ such that

$$|||u - u_2|||^2 \leq \Lambda |||u - u_1|||^2 - |||u_2 - u_1|||^2.$$

2) Error estimator as upper bound on error (§2.3.2): There exists $C_1 > 0$ such that

$$|||u - u_k|||^2 \leq C_1 \eta_k^2(u_k, \mathscr{T}_k), \quad k = 1, 2.$$

3) Estimator reduction (§2.3.4): For $\mathscr{M}$ the marked set that takes refinement $\mathscr{T}_1 \to \mathscr{T}_2$, for positive constants $\lambda < 1$ and Λ_1 and any $\delta > 0$

$$\eta_2^2(v_2, \mathscr{T}_2) \leq (1+\delta)\{\eta_1^2(v_1, \mathscr{T}_1) - \lambda \eta_1^2(v_1, \mathscr{M}\} + (1+\delta^{-1})\Lambda_1 \eta_0^2 |||v_2 - v_1|||.$$

The marking strategy used is the original Dörfler strategy; elements are marked for refinement based on indicators alone. The marked set $\mathscr{M}$ must satisfy

$$\sum_{T\in\mathscr{M}} \eta_k^2(u_k,T) \geq \theta^2\eta_k^2(u_k,\mathscr{T}_k).$$

In the goal-oriented method, a second marked set is chosen based on an error indicator for the dual problem associated with the given goal functional, and the union of the two marked sets is then used for refinement.

A main advantage of the approach in [4] is that it does not require an interior node property. This allows us to establish the necessary results for contraction without taking full refinements of the mesh at each iteration. This improvement follows from the use of the local perturbation estimate or local Lipschitz property rather than the estimator as lower bound on error. We use the standard lower bound estimate as found in [9] for optimality arguments in the second part of the paper concerning quasi-optimality of the method.

There are three main notions of error used throughout this paper. The energy error $|||u-u_k|||$, the quasi-error and the total-error. The *energy error* is defined by the symmetric part of the bilinear form that arises from the given differential operator in (2.1.3). The *quasi-error* is the l_2 sum of the energy-error and scaled error estimator

$$Q_k(u_k,\mathscr{T}_k) := (|||u-u_k|||^2 + \gamma\eta_k^2)^{1/2},$$

and this is the quantity that is reduced at each iteration of the algorithm. In §2.3 the quasi-error is shown to satisfy

$$|||u-u_{k+1}|||^2 + \gamma\eta_{k+1}^2 \leq \alpha^2\left(|||u-u_k|||^2 + \gamma\eta_k^2\right),\ \alpha < 1.$$

The *total error* includes the oscillation term rather than the estimator

$$E_k(u_k,\mathscr{T}_k) := (|||u-u_k|||^2 + \operatorname{osc}_k^2)^{1/2}.$$

The oscillation term captures the higher-frequency oscillations in the residual missed by the averaging of the finite element method. While the quasi-error is the focus of the contraction arguments, it is the total error that will be critical to complexity analysis. Therefore, we will need to establish various preliminary results for both types of error.

The quasi-optimality of the goal oriented method in §2.4 is developed with respect to the total error which is shown to satisfy Cea's lemma. The cardinality result

$$\begin{aligned}\#\mathscr{T}_k - \#\mathscr{T}_0 \leq S_\theta \Bigg\{ & M_p \left(1+\frac{\gamma_p}{c_2}\right)^{1/2s} Q_k^{-1/s}(u_k, \mathscr{T}_k) \\ & + M_d \left(1+\frac{\gamma_d}{c_2}\right)^{1/2t} Q_k^{-1/t}(z_k, \mathscr{T}_k) \Bigg\}\end{aligned}$$

bounds the growth of the adaptive mesh with respect to the quasi-error of both problems. An equivalence between the quasi-error and total error is established in §2.4.

A final brief comment is in order concerning the notation used here compared to that in [4] and the related literature. In [4], the number of times each marked element is refined is denoted b. In this article, each marked element is refined once. Therefore, b will be reserved for the convection term in the nonsymmetric problem. The constant C will denote a generic but global constant that may depend on the data and the condition of the initial mesh $\mathscr{T}_0$, and may change from step to step.

Outline of the paper. The remainder of the paper is structured as follows. In §2.2, we first describe the problem class and review some standard facts concerning conforming finite element discretization and error-estimate-driven adaptive finite element methods (AFEM). In §2.2.3, we then describe a goal-oriented variation of the standard approach to AFEM (GOAFEM). Following the recent work of Mommer and Stevenson for symmetric problems, in §2.3 we establish contraction of goal-oriented AFEM. We also then show convergence in §2.3.6 in the sense of the goal function. Our analysis approach is signficantly different, combining the recent contraction frameworks developed by Cascon et. al [4], Nochetto, Siebert, and Veeser [11], and by Holst, Tsogtgerel, and Zhu [8]. In §2.4, we consider complexity questions, and establish quasi-optimal cardinality of GOAFEM. We recap the results in §2.5, and point out some remaining open problems.

2.2 Problem class, discretization, goal-oriented AFEM

2.2.1 Problem class, weak formulation, spaces and norms

Consider the nonsymmetric problem (2.1.3), where as in (2.1.4) we have

$$a(u,v) = \langle A\nabla u, \nabla v\rangle + \langle b \cdot \nabla u, v\rangle + \langle cu, v\rangle.$$

Here we have introduced the notation $\langle \cdot, \cdot \rangle$ for the L_2 inner-product over $\Omega \subset \mathbb{R}^d$. The adjoint or dual problem is: Find $z \in H_0^1(\Omega)$ such that

$$a^*(z,v) = g(v) \ \text{ for all } v \in H_0^1(\Omega) \tag{2.2.1}$$

where $a^*(\cdot\,,\cdot)$ is the formal adjoint of $a(\cdot\,,\cdot)$, and where the functional is defined through

$$g(u) = \int_\Omega gu\, dx, \tag{2.2.2}$$

for some given $g \in L_2(\Omega)$. We will make the following assumptions on the data:

Assumption 2.2.1 (Problem data)**.** *The problem data $D = (A,b,c,f)$ and dual problem data $D^* = (A,-b,c,g)$ satisfy*

1) $A : \overline{\Omega} \to \mathbb{R}^{d\times d}$, Lipschitz, and a.e. symmetric positive-definite:

$$ess\ inf_{x\in\Omega}\lambda_{min}(A(x)) = \mu_0 > 0, \tag{2.2.3}$$

$$ess\ sup_{x\in\Omega}\lambda_{max}(A(x)) = \mu_1 < \infty. \tag{2.2.4}$$

2) $b : \overline{\Omega} \to \mathbb{R}^d$, with $b_k \in L_\infty(\Omega)$, and b divergence-free.

3) $c : \overline{\Omega} \to \mathbb{R}$, with $c \in L_\infty(\Omega)$, and $c(x) \geq 0$ for all $x \in \Omega$.

4) $f, g \in L_2(\Omega)$.

The native norm is the Sobolev H^1 norm given by

$$\|v\|_{H^1}^2 = \langle \nabla v, \nabla v\rangle + \langle v, v\rangle. \tag{2.2.5}$$

The L_p norm of a vector valued function v over domain ω is defined here as the l_2 norm of the $L_p(\omega)$ norm of each component

$$\|v\|_{L_p(\omega)} = \left(\sum_{j=1}^{d} \left(\int_{\omega} v_j^p \right)^{2/p} \right)^{1/2}, \; p = 1, 2, \ldots$$
$$\|v\|_{L_\infty(\omega)} = \left(\sum_{j=1}^{d} \left(\operatorname*{ess\,sup}_{\omega} v_j \right)^{2} \right)^{1/2}. \tag{2.2.6}$$

Similarly, the L_p norm of a matrix valued function M over domain ω is defined as the Frobenius norm of the $L_p(\omega)$ norm of each component

$$\|M\|_{L_p(\omega)} = \left(\sum_{i,j=1}^{d} \left(\int_{\omega} M_{ij}^p \right)^{2/p} \right)^{1/2}, \quad p = 1, 2, \ldots$$
$$\|M\|_{L_\infty(\omega)} = \left(\sum_{ij=1}^{d} \left(\operatorname*{ess\,sup}_{\omega} M_{ij} \right)^{2} \right)^{1/2}. \tag{2.2.7}$$

We note that one could employ other equivalent discrete l_p norms in the definitions (2.2.6) and (2.2.7), however this choice simplifies the analysis.

Continuity of $a(\cdot,\cdot)$ follows from the Hölder inequality, and bounding the L_2 norm of the function and its gradient by the H^1 norm

$$a(u,v) \leq (\mu_1 + \|b\|_{L_\infty} + \|c\|_{L_\infty}) \|u\|_{H^1} \|v\|_{H^1} = M_c \|u\|_{H^1} \|v\|_{H^1}. \tag{2.2.8}$$

Coercivity follows from the Poincaré inequality with constant C_Ω and the divergence-free condition

$$a(v,v) \geq \mu_0 |v|_{H^1}^2 \geq C_\Omega \mu_0 \|v\|_{H^1}^2 = m_{\mathscr{E}}^2 \|v\|_{H^1}^2, \tag{2.2.9}$$

where the coercivity constant $m_{\mathscr{E}}^2 := C_\Omega \mu_o$. Continuity and coercivity imply existence and uniqueness of the solution by the Lax-Milgram Theorem [7]. The adjoint operator $a^*(\ ,\ )$ is given by

$$a^*(v,u) := a(u,v), \qquad u,v \in H_0^1(\Omega).$$

Integration by parts on the convection term and the divergence-free condition imply

$$a^*(z,v) := \langle A\nabla z, \nabla v\rangle - \langle b\cdot\nabla z, v\rangle + \langle cz, v\rangle. \tag{2.2.10}$$

Define the energy semi-norm by

$$|||v|||^2 := a(v,v). \tag{2.2.11}$$

Non-negativity follows directly from the coercivity estimate (2.2.9)

$$|||v|||^2 \geq m_{\mathscr{E}}^2 \|v\|_{H^1}^2, \tag{2.2.12}$$

which establishes the energy semi-norm as a norm. Putting this together with the reverse inequality

$$|||v|||^2 \leq \mu_1 |\nabla v|_{L_2}^2 + \|c\|_{L_\infty} \|v\|_{L_2}^2 \implies |||v||| \leq M_{\mathscr{E}} \|v\|_{H^1}, \tag{2.2.13}$$

establishes the equivalence between the native and energy norms with the constant $M_{\mathscr{E}} = (\mu_1 + \|c\|_{L_\infty})^{1/2}$.

2.2.2 Finite element approximation

We employ a standard conforming piecewise polynomial finite element approximation below.

Assumption 2.2.2 (Finite element mesh)**.** *We make the following assumptions on the underlying simplex mesh:*

1) *The initial mesh $\mathscr{T}_0$ is conforming.*

2) *The mesh is refined by newest vertex bisection [2], [10] at each iteration.*

3) *The initial mesh $\mathscr{T}_0$ is sufficiently fine. In particular, it satisfies (2.3.6).*

Based on assumptions 2.2.2 we have the following mesh constants.

1) Define

$$h_{\mathscr{T}} := \max_{T\in\mathscr{T}} h_T, \quad \text{where } h_T = |T|^{1/d}. \tag{2.2.14}$$

In particular, h_0 is the initial mesh diameter.

2) Define the mesh constant $\gamma_N = 2\gamma_r$ where

$$\gamma_r = \frac{h_0}{h_{min}} \quad \text{and } h_{min} = \min_{T \in \mathscr{T}_0} h_T$$

then for any two elements $T, \tilde{T}$ in the same generation

$$h_T \leq \gamma_r h_{\tilde{T}}$$

and as neighboring elements may differ by at most one generation for any two neighboring elements T and T'

$$h_T \leq 2\gamma_r h_{T'} = \gamma_N h_{T'}. \tag{2.2.15}$$

3) The minimal angle condition satisfied by newest vertex bisection implies the meshsize h_T is comparable to h_σ, the size of any true-hyperface σ of T. In particular, there is a constant $\bar{\gamma}$

$$\frac{h_\sigma}{h_T} \leq \bar{\gamma}^2 \text{ for all } T. \tag{2.2.16}$$

Let $\mathbb{T}$ the set of conforming meshes derived from the initial mesh $\mathscr{T}_0$. Define $\mathbb{T}_N \subset \mathbb{T}$ by

$$\mathbb{T}_N = \{\mathscr{T} \in \mathbb{T} \mid \#\mathscr{T} - \#\mathscr{T}_0 \leq N\}.$$

For a conforming mesh $\mathscr{T}_1$ with a conforming refinement $\mathscr{T}_2$ we say $\mathscr{T}_2 \geq \mathscr{T}_1$. The set of refined elements is given by

$$\mathscr{R}_{1\to 2} := \mathscr{R}_{\mathscr{T}_1 \to \mathscr{T}_2} := \mathscr{T}_1 \setminus (\mathscr{T}_2 \cap \mathscr{T}_1). \tag{2.2.17}$$

An *overlay* of two meshes $\mathscr{T}_1 \geq \mathscr{T}_0$ and $\mathscr{T}_2 \geq \mathscr{T}_0$ where $\mathscr{T}_2$ is not generally a refinement of $\mathscr{T}_1$ is given by

$$\mathscr{T}_1 \oplus \mathscr{T}_2 := \{T \in \mathscr{T}_1 \big| T \subseteq T' \text{ for some } T' \in \mathscr{T}_2\} \cup \{T \in \mathscr{T}_2 \big| T \subseteq T' \text{ for some } T' \in \mathscr{T}_1\} \tag{2.2.18}$$

and is itself conforming. Define the finite element space

$$\mathbb{V}_{\mathscr{T}} := H_0^1(\Omega) \cap \prod_{T \in \mathscr{T}} \mathbb{P}_n(T) \quad \text{ and } \mathbb{V}_k := \mathbb{V}_{\mathscr{T}_k}. \tag{2.2.19}$$

For subsets $\omega \subseteq \mathscr{T}$,

$$\mathbb{V}_{\mathscr{T}}(\omega) := H_0^1(\Omega) \cap \prod_{T \in \omega} \mathbb{P}_n(T), \tag{2.2.20}$$

where $\mathbb{P}_n(T)$ is the space of polynomials degree degree n over T. Denote the patch about $T \in \mathscr{T}$

$$\omega_T := T \cup \{T' \in \mathscr{T} \mid T \cap T' \text{ is a true-hyperface of } T\}. \tag{2.2.21}$$

For a d-simplex T, an true-hyperface is a $d-1$ dimensional face of T, *e.g.,* a face in 3D or an edge in 2D. Define the discrete primal problem: Find $u_k \in \mathbb{V}_k$ such that

$$a(u_k, v_k) = f(v_k), \; v_k \in \mathbb{V}_k, \tag{2.2.22}$$

and the discrete dual problem

$$a^*(z_k, v_k) = g(v_k), \; v_k \in \mathbb{V}_k. \tag{2.2.23}$$

2.2.3 Goal oriented AFEM (GOAFEM)

As in [10] the goal oriented adaptive finite element method (GOAFEM) is based on the standard AFEM algorithm:

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}\,.$$

In the goal oriented method, one enforces contraction of the quasi-error in both the primal problem and an associated dual problem. As shown in section §2.3.6, the error in the goal-function satisfies the bound

$$|g(u)-g(u_k)| = |a(u-u_k, z-z_k)| \leq 2|||u-u_k||| \, |||z-z_k|||.$$

This motivates driving down the energy-error in both the primal and dual problems at each iteration. As noted in [4] the residual-based error estimator does not exhibit monotone behavior in general, although it is monotone non-increasing with respect to nested mesh refinement when applied to the same (coarse) function. The quasi-error is shown to contract for each problem for which mesh refinement satisfies the Dörfler property. However, refining the mesh with respect to the primal problem does not guarantee the quasi-error in the dual problem will be non-increasing, and vice-versa. As such, the procedures SOLVE and ESTIMATE are performed for each of the primal and dual problems. The marked set is taken to be the union of marked sets from the primal and dual problems, each chosen to satisfy the Dörfler property. This method produces a sequence of refinements for which both the error in the primal and dual problems contract at each step.

Procedure SOLVE. The contraction result supposes the exact Galerkin solution is found on each mesh refinement. In practice a linear-time iterative method is employed so that the Galerkin solution is found up to a given tolerance.

Procedure ESTIMATE. The estimation of the error on each element is determined by a standard residual-based estimator. The residuals over element interiors and jump-residuals over the boundaries are based on the *local strong forms* of the elliptic operator and its adjoint as follows.

$$\mathscr{L}(v) = \nabla \cdot (A\nabla v) - b \cdot \nabla v - cv; \quad \mathscr{L}^*(v) = \nabla \cdot (A\nabla v) + b \cdot \nabla v - cv. \tag{2.2.24}$$

The *residuals* for the primal and dual problems using the sign convention in [4] are:

$$R(v) := f + \mathscr{L}(v); \quad R^*(v) := g + \mathscr{L}^*(v), \; v \in \mathbb{V}_{\mathscr{T}}. \tag{2.2.25}$$

While the primal and dual solutions u and z of (2.1.3) and (2.2.1) respectively satisfy

$$f(z) = a(u,z) = a^*(z,u) = g(u)$$

the residuals for the primal and dual problems are in general different. The *jump residual* for the primal and dual problems is

$$J_T(v) := [\![[A\nabla v]\cdot n]\!]_{\partial T} \tag{2.2.26}$$

where *jump operator* $[\![\; \cdot \;]\!]$ is given by

$$[\![\phi]\!]_{\partial T} := \lim_{t\to 0} \phi(x+tn) - \phi(x-tn) \tag{2.2.27}$$

and n is taken to be the appropriate outward normal defined piecewise on ∂T. On boundary edges σ_b we have

$$[\![[A\nabla v]\cdot n]\!]_{\sigma_b} \equiv 0$$

so that $[\![[A\nabla v]\cdot n]\!]_{\partial T} = [\![[A\nabla v]\cdot n]\!]_{\partial T\cap\Omega}$. For clarity, we will also employ the notation

$$R_T(v) := R(v)\big|_T,\; v \in \mathbb{V}_{\mathscr{T}},$$

and similarly for the other strong form operators. The error indicator is given as

$$\eta^p_{\mathscr{T}}(v,T) := h^p_T \|R(v)\|^p_{L_2(T)} + h^{p/2}_T \|J_T(v)\|^p_{L_2(\partial T)}, \quad v \in \mathbb{V}_{\mathscr{T}}. \tag{2.2.28}$$

The dual error-indicator is then given by

$$\zeta^p_{\mathscr{T}}(w,T) := h^p_T \|R^*(w)\|^p_{L_2(T)} + h^{p/2}_T \|J_T(w)\|^p_{L_2(\partial T)}, \quad w \in \mathbb{V}_{\mathscr{T}}. \tag{2.2.29}$$

The error estimators are given by the l_p sum of error indicators over elements in the space where $p = 1$ or 2.

$$\eta^p_{\mathscr{T}}(v) := \sum_{T\in\mathscr{T}} \eta^p_{\mathscr{T}}(v,T), \quad v \in \mathbb{V}_{\mathscr{T}}. \tag{2.2.30}$$

The dual energy estimator is:

$$\zeta^p_{\mathscr{T}}(w) := \sum_{T\in\mathscr{T}} \zeta^p_{\mathscr{T}}(w), \quad w \in \mathbb{V}_{\mathscr{T}}. \tag{2.2.31}$$

The contraction results for the quasi-error presented below will be shown to hold for $p = 1,2$ where the error estimator and oscillation are defined in terms of the l_p norm. While complexity results are shown only for $p = 2$, the contraction results for $p = 1$ are useful for nonlinear problems; see [8].

For analyzing oscillation, for $v \in \mathbb{V}_{\mathscr{T}}$ let Π^2_m the orthogonal projector defined by the best L_2 approximation in $\mathbb{P}_m$ over mesh $\mathscr{T}$ and $P^2_m = I - \Pi^2_m$. Define now the oscillation on the elements $T \in \mathscr{T}$ for the primal problem by

$$\operatorname{osc}_{\mathscr{T}}(v,T) := h_T \|P^2_{2n-2}R(v)\|_{L_2(T)} \tag{2.2.32}$$

and analogously for the dual problem. For subsets $\omega \subseteq \mathscr{T}$ set

$$\operatorname{osc}^p_{\mathscr{T}}(v,\omega) := \sum_{T\in\omega} \operatorname{osc}^p_{\mathscr{T}}(v,T). \tag{2.2.33}$$

The data estimator and data oscillation, identical for both the primal and dual problems, are given by

$$\eta^p_{\mathscr{T}}(D,T) := h^p_T \left(\|\operatorname{div} A\|^p_{L_\infty(T)} + h_T^{-p}\|A\|^p_{L_\infty(\omega_T)} + \|c\|^p_{L_\infty(T)} + \|b\|^p_{L_\infty(T)} \right), \tag{2.2.34}$$

$$\begin{aligned}\operatorname{osc}^p_{\mathscr{T}}(D,T) := h^p_T \Big(&\|P^\infty_{n-1}\operatorname{div} A\|^p_{L_\infty(T)} + h_T^{-p}\|P^\infty_n A\|^p_{L_\infty(T)} \\ &+ h^p_T\|P^\infty_{n-2}c\|^p_{L_\infty(T)} + \|P^\infty_{2n-2}c\|^p_{L_\infty(T)} + \|P^\infty_{n-1}b\|^p_{L_\infty(T)} \Big).\end{aligned} \tag{2.2.35}$$

The data estimator and oscillation over the mesh $\mathscr{T}$ or a subset $\omega \subset \mathscr{T}$ are given by the maximum data estimator (oscillation) over elements in the mesh or subset: For $\omega \subseteq \mathscr{T}$

$$\eta_{\mathscr{T}}(D,\omega) = \max_{T\in\omega} \eta_{\mathscr{T}}(D,T) \text{ and } \operatorname{osc}_{\mathscr{T}}(D,\omega) = \max_{T\in\omega} \operatorname{osc}_{\mathscr{T}}(D,T).$$

The data estimator and data oscillation on the initial mesh

$$\eta_0 := \eta_{\mathscr{T}_0}(D,\mathscr{T}_0), \text{ and } \operatorname{osc}_0 := \operatorname{osc}_{\mathscr{T}_0}(D,\mathscr{T}_0).$$

As the grid is refined, the data estimator and data oscillation terms satisfy the monotonicity property [4] for refinements $\mathscr{T}_2 \geq \mathscr{T}_1$

$$\eta_2(D,\mathscr{T}_2) \leq \eta_1(D,\mathscr{T}_1) \text{ and } \operatorname{osc}_2(D,\mathscr{T}_2) \leq \operatorname{osc}_1(D,\mathscr{T}_1). \tag{2.2.36}$$

Procedure MARK. The Dörfler marking strategy for the goal-oriented problem is based on the following steps as in [10]:

1) Given $\theta \in (0,1)$, mark sets for each of the primal and dual problems:

   - Mark a set $\mathscr{M}_p \subset \mathscr{T}_k$ such that,

$$\sum_{T \in \mathscr{M}_p} \eta_k^2(u_k,T) \geq \theta^2 \eta_k^2(u_k,\mathscr{T}_k) \tag{2.2.37}$$

   - Mark a set $\mathscr{M}_d \subset \mathscr{T}_k$ such that,

$$\sum_{T \in \mathscr{M}_d} \zeta_k^2(z_k,T) \geq \theta^2 \zeta_k^2(z_k,\mathscr{T}_k) \tag{2.2.38}$$

2) Let $\mathscr{M} = \mathscr{M}_p \cup \mathscr{M}_d$ the union of sets found for the primal and dual problems respectively.

The set $\mathscr{M}$ differs from that in [10], where the set of lesser cardinality between $\mathscr{M}_p$ and $\mathscr{M}_d$ is used. In the case of the nonsymmetric problem the error reduced at each iteration is the quasi-error rather than the energy error as in the symmetric problem [10]. This error for each problem is guaranteed to contract based on the refinement satisfying the Dörfler property. As such, refining the mesh with respect to one problem does not guarantee the quasi-error in the other problem is nonincreasing. Sets $\mathscr{M}_p$ and $\mathscr{M}_d$ with optimal cardinality (up to a factor of 2) can be chosen in linear time by binning the elements rather than performing a full sort [10].

Procedure REFINE. The refinement (including the completion) is performed according to newest vertex bisection [2]. The complexity and other properties of this procedure are now well-understood, and will simply be exploited here.

2.3 Contraction and convergence theorems

The key elements of the main contraction argument constructed below are quasi-orthogonality 2.3.1, error estimator as upper-bound on energy-norm error 2.3.2 and estimator reduction 2.3.4. Estimator-reduction is shown via the local-perturbation estimate 2.3.3. The local perturbation of the oscillation is presented here and used in §2.4. Mesh refinements $\mathscr{T}_1$ and $\mathscr{T}_2$ (respectively $\mathscr{T}_j$) are assumed conforming, and u_j is assumed the Galerkin solution on refinement $\mathscr{T}_j$. The following results hold for both the primal and dual problems which differ by the sign of the convection term; therefore, they are established here only for the primal problem.

2.3.1 Quasi-orthogonality

Orthogonality in the energy-norm $|||u-u_2|||^2 = |||u-u_1|||^2 - |||u_2-u_1|||^2$ does not generally hold in the nonsymmetric problem. We use the weaker quasi-orthogonality result to establish contraction of AFEM (GOAFEM). The following is a variation on Lemma 2.1 in [9] (see also [8]).

Lemma 2.3.1 (Quasi-orthogonality)**.** *Let the problem data satisfy Assumption 2.2.1 and the mesh satisfy conditions (1) and (2) of Assumption 2.2.2. Let $\mathscr{T}_1, \mathscr{T}_2 \in \mathbb{T}$ with $\mathscr{T}_2 \geq \mathscr{T}_1$. Let $u_k \in \mathbb{V}_k$ the solution to* (2.2.22)*, $k = 1,2$. There exists a constant $C_* > 0$ depending on the problem data D and initial mesh $\mathscr{T}_0$, and a number $0 < s \leq 1$ dictated only by the angles of $\partial\Omega$, such that if the meshsize h_0 of the initial mesh satisfies $\bar{\Lambda} := C_* h_0^s \|b\|_{L_\infty} \mu_0^{-1/2} < 1$, then*

$$|||u-u_2|||^2 \leq \Lambda |||u-u_1|||^2 - |||u_2-u_1|||^2, \tag{2.3.1}$$

where

$$\Lambda := (1 - C_* h_0^s \|b\|_{L_\infty} \mu_0^{-1/2})^{-1}.$$

Equality holds (usual orthogonality) when $b = 0$ in Ω, in which case the problem is symmetric.

Proof. The proof follows close that of Lemma 2.1 in [9]. Let

$$e_2 := u - u_2, \quad e_1 := u - u_1, \quad \text{and } \varepsilon_1 := u_2 - u_1.$$

By Galerkin orthogonality

$$|||e_1|||^2 = a(e_1, e_1) = |||e_2|||^2 + |||\varepsilon_1|||^2 + a(\varepsilon_1, e_2). \tag{2.3.2}$$

Rearranging and applying the divergence-free condition on the convection term

$$|||e_2|||^2 = |||e_1|||^2 - |||\varepsilon_1|||^2 - 2\langle b \cdot \nabla \varepsilon_1, e_2 \rangle.$$

Applying Hölder's inequality and coercivity (2.2.9) $|\varepsilon_1|_{H^1} \leq \mu_0^{-1/2} |||\varepsilon_1|||$ followed by Young's inequality with constant δ to be determined,

$$-2\langle b \cdot \nabla \varepsilon_1, e_2 \rangle \leq \delta \|e_2\|_{L_2}^2 + \frac{\|b\|_{L_\infty}^2}{\delta \mu_0} |||\varepsilon_1|||^2. \tag{2.3.3}$$

By a duality argument for some $C_* > 0$ assuming $u \in H^{1+s}(\Omega)$ for some $0 < s \leq 1$ depending on the angles of $\partial\Omega$

$$\|e_2\|_{L_2} \leq C_* h_0^s |||e_2|||. \tag{2.3.4}$$

The details of this argument as described in the appendix §2.6 may also be found in [1] and [5]. Applying (2.3.4) and (2.3.3) to (2.3.2),

$$(1 - \delta C_*^2 h_0^{2s}) |||u - u_2|||^2 \leq |||u - u_1|||^2 - \left(1 - \frac{\|b\|_{L_\infty}^2}{\delta \mu_0}\right) |||u_1 - u_2|||^2. \tag{2.3.5}$$

Choose δ to equate coefficients

$$\delta C_*^2 h_0^{2s} = \frac{\|b\|_{L_\infty}^2}{\delta \mu_0} \implies \delta = \frac{\|b\|_{L_\infty}}{C_* h_0^s \sqrt{\mu_0}},$$

then

$$|||u-u_2|||^2 \leq \left(1-\|b\|_{L_\infty} C_* h_0^s \mu_0^{-1/2}\right)^{-1} |||u-u_1|||^2 - |||u_1-u_2|||^2.$$

Assuming the initial mesh as characterized by h_0 satisfies

$$\bar{\Lambda} = \|b\|_{L_\infty} C_* h_0^s \mu_0^{-1/2} < 1, \tag{2.3.6}$$

the quasi-orthogonality result holds. □

Note that by (2.3.2) we also have

$$|||\varepsilon_1|||^2 = |||e_1|||^2 - |||e_2|||^2 - 2\langle b\cdot\nabla e_2, \varepsilon_1\rangle. \tag{2.3.7}$$

Similarly to (2.3.3)

$$-2\langle b\cdot\nabla e_2, \varepsilon_1\rangle \geq -2|\langle b\cdot\nabla e_2, \varepsilon_1\rangle| \geq -\delta\|\varepsilon_1\|_{L_2}^2 - \frac{\|b\|_{L_\infty}^2}{\delta\mu_0}|||e_2|||^2, \tag{2.3.8}$$

which under the same assumptions yields the estimate

$$|||u_2-u_1|||^2 \geq (1+\bar{\Lambda})^{-1}|||u-u_1|||^2 - |||u-u_2|||^2, \tag{2.3.9}$$

where $\bar{\Lambda} < 1 \implies (1+\bar{\Lambda})^{-1} > 1/2$.

2.3.2 Error estimator as global upper-bound

We now recall the property that the error estimator is a global upper bound on the error. The proof is fairly standard; see e.g. [10] (Proposition 4.1), [9] (3.6), and [8].

Lemma 2.3.2 (Error estimator as global upper-bound)**.** *Let the problem data satisfy Assumption 2.2.1 and the mesh satisfy conditions (1) and (2) of Assumption 2.2.2. Let $\mathscr{T}_1, \mathscr{T}_2 \in \mathbb{T}$ with $\mathscr{T}_2 \geq \mathscr{T}_1$. Let $u_k \in \mathbb{V}_k$ the solution to* (2.2.22)*, $k = 1,2$ and u the solution to* (2.1.3)*. Let*

$$G = G(\mathscr{T}_2, \mathscr{T}_1) := \{T \subset \mathscr{T}_1 \mid T\cap\tilde{T} \neq \emptyset \textit{ for some } \tilde{T}\in\mathscr{T}_1, \tilde{T}\notin\mathscr{T}_2\}.$$

Then for global constant C_1 depending on the problem data D and initial mesh $\mathscr{T}_0$

$$|||u_2 - u_1||| \leq C_1 \eta_1(u_1, G) \tag{2.3.10}$$

and in particular

$$|||u - u_1||| \leq C_1 \eta_1(u_1, \mathscr{T}_1). \tag{2.3.11}$$

2.3.3 Local perturbation

The local perturbation property established in [4], analogous to the local Lipshitz property in [8], is a key step in establishing the contraction result. This is a minor variation on Proposition 3.3 in [4] which deals with a symmetric problem. Here, we include a convection term in the estimate. In particular, (2.3.12) shows that the difference in the error indicators over an element T between two functions in a given finite element space may be bounded by a fixed factor of the native norm over the patch ω_T of the difference in functions. In contrast with the analogous result in [4] the estimate (2.3.13) involves a fixed factor of the native norm over an individual element rather than a patch as by the continuity of A the oscillation term does not involve the jump residual.

We include the proof of (2.3.12) for completeness. The proof of (2.3.13) may be found in [4] with the final result inferred by the absence of the jump residual in the oscillation term.

Lemma 2.3.3 (Local perturbation)**.** *Let the problem data satisfy Assumption 2.2.1 and the mesh satisfy condition (1) of Assumption 2.2.2. Let $\mathscr{T} \in \mathbb{T}$. For all $T \in \mathscr{T}$ and for any $v, w \in \mathbb{V}_{\mathscr{T}}$*

$$\eta_{\mathscr{T}}(v,T) \leq \eta_{\mathscr{T}}(w,T) + \bar{\Lambda}_1 \eta_{\mathscr{T}}(D,T) \|v - w\|_{H^1(\omega_T)} \tag{2.3.12}$$

$$\operatorname{osc}_{\mathscr{T}}(v,T) \leq \operatorname{osc}_{\mathscr{T}}(w,T) + \bar{\Lambda}_2 \operatorname{osc}_{\mathscr{T}}(D,T) \|v - w\|_{H^1(T)} \tag{2.3.13}$$

where recalling (2.2.21) *ω_T is the union of T with elements in $\mathscr{T}$ sharing a true-hyperface with T. The constants $\bar{\Lambda}_1, \bar{\Lambda}_2 > 0$ depend on the initial mesh $\mathscr{T}_0$, the dimension d and the polynomial degree n.*

Proof of (2.3.12). From (2.2.28)

$$\eta^p_{\mathscr{T}}(v,T) := h^p_T \|R(v)\|^p_{L_2(T)} + h^{p/2}_T \|J_T(v)\|^p_{L_2(\partial T)}, \quad v \in \mathbb{V}_{\mathscr{T}}. \tag{2.3.14}$$

Denote $\eta_{\mathscr{T}}(v,T)$ by $\eta(v,T)$. Set $e = v - w$. By linearity

$$R(v) = R(w+e) = f + \mathscr{L}(w+e) = f + \mathscr{L}(w) + \mathscr{L}(e) = R(w) + \mathscr{L}(e)$$

and

$$J(v) = J(w+e) = J(w) + J(e).$$

For $p = 1$ by the triangle inequality

$$\begin{aligned}\eta(v,T) &= h_T \|R(w) + \mathscr{L}(e)\|_{L_2(T)} + h^{1/2}_T \|J(w) + J(e)\|_{L_2(\partial T)} \\ &\le \eta(w,T) + h_T \|\mathscr{L}(e)\|_{L_2(T)} + h^{1/2}_T \|J(e)\|_{L_2(\partial T)}.\end{aligned}$$

For $p = 2$ using the generalized triangle-inequality

$$\sqrt{(a+b)^2 + (c+d)^2} \le \sqrt{a^2 + c^2} + b + d, \qquad \text{for } a,b,c,d > 0 \tag{2.3.15}$$

we have

$$\begin{aligned}\eta(v,T) &= \Big(h^2_T \|R(w) + \mathscr{L}(e)\|^2_{L_2(T)} + h_T \|J(w) + J(e)\|^2_{L_2(\partial T)}\Big)^{1/2} \\ &\le \eta(w,T) + h_T \|\mathscr{L}(e)\|_{L_2(T)} + h^{1/2}_T \|J(e)\|_{L_2(\partial T)}.\end{aligned}$$

Consider the second term on the RHS $h_T \|\mathscr{L}(e)\|_{L_2(T)}$. By definition (2.2.24) of $\mathscr{L}(\,\cdot\,)$, the product rule applied to the diffusion term and the triangle-inequality

$$\|\mathscr{L}(e)\|_{L_2(T)} \le \|\mathrm{div} A \cdot \nabla e\|_{L_2(T)} + \|A : D^2 e\|_{L_2(T)} + \|ce\|_{L_2(T)} + \|b \cdot \nabla e\|_{L_2(T)}$$

where D^2e is the Hessian of e. Consider each term. The first diffusion term

$$\|\text{div}A \cdot \nabla e\|_{L_2(T)} \leq \|\text{div}A\|_{L_\infty(T)} \|\nabla e\|_{L_2(T)} \tag{2.3.16}$$

by the inequality

$$\|v \cdot z\|_{L_2(T)} \leq \|v\|_{L_\infty(T)} \|z\|_{L_2(T)}, \quad v \in L_\infty(T),\ z \in L_2(T). \tag{2.3.17}$$

Applying (2.3.17) and inverse-estimate [3] to the second diffusion term

$$\begin{aligned} \|A : D^2e\|_{L_2(T)} &\leq \|A\|_{L_\infty(T)} \|D^2e\|_{L_2(T)} \\ &\leq C_I h_T^{-1} \|A\|_{L_\infty(T)} \|\nabla e\|_{L_2(T)}. \end{aligned} \tag{2.3.18}$$

For the reaction term

$$\|ce\|_{L_2(T)} \leq \|c\|_{L_\infty(T)} \|e\|_{L_2(T)}. \tag{2.3.19}$$

For the convection term applying (2.3.17)

$$\|b \cdot \nabla e\|_{L_2(T)} \leq \|b\|_{L_\infty(T)} \|\nabla e\|_{L_2(T)}. \tag{2.3.20}$$

Consider the the jump-residual term $\|J(e)\|_{L_2(\partial T)}$. For each interior true-hyperface $\sigma = T \cap T',\ T, T' \in \mathscr{T}$ by (2.2.27)

$$\begin{aligned} J(e)\big|_\sigma &:= \lim_{t \to 0^+} (A\nabla e)(x + tn_\sigma) - \lim_{t \to 0^-} (A\nabla e)(x - tn_\sigma) \\ &= n_\sigma \cdot (A\nabla e)\big|_T - n_\sigma \cdot (A\nabla e)\big|_{T'} \end{aligned} \tag{2.3.21}$$

where $(A\nabla e)\big|_T$ is understood to refer to the product of the limiting value of $A\nabla e$ as the element boundary is approached from the interior of T. By the triangle-inequality

$$\|J(e)\|_{L_2(\sigma)} \leq \|n_\sigma \cdot (A\nabla e)\big|_T\|_{L_2(\sigma)} + \|n_\sigma \cdot (A\nabla e)\big|_{T'}\|_{L_2(\sigma)}.$$

By bounds for the inner-product with a unit normal and a matrix-vector product

$$\|\phi \cdot n\|_{L_2(\sigma)} \leq \|\phi\|_{L_2(\sigma)}, \quad \phi \in L_2(\sigma), \tag{2.3.22}$$

$$\|M\phi\|_{L_2(T)} \leq \|M\|_{L_\infty(T)}\|\phi\|_{L_2(T)}, \quad M \in L_\infty(T),\ \phi \in L_2(T) \tag{2.3.23}$$

obtain

$$\|n_\sigma \cdot (A\nabla e)\big|_T\|_{L_2(\sigma)} \leq \|(A\nabla e)\big|_T\|_{L_2(\sigma)} \leq \|A\big|_T\|_{L_\infty(\sigma)}\|\nabla e\big|_T\|_{L_2(\sigma)}. \tag{2.3.24}$$

Applying the trace theorem and an inverse inequality to $\|\nabla e\big|_T\|_{L_2(\sigma)}$ via the inequality

$$\|\phi\|_{L_2(\sigma)} \leq Ch_T^{-1/2}\|\phi\|_{L_2(T)}, \quad \phi \in L_2(T) \tag{2.3.25}$$

we have

$$\|\nabla e\big|_T\|_{L_2(\sigma)} \leq C_T(\bar{\gamma})^{d-1}h_T^{-1/2}\|\nabla e\|_{L_2(T)}. \tag{2.3.26}$$

By the Lipschitz property of A

$$\|A\big|_T\|_{L_\infty(\sigma)} = \|A\|_{L_\infty(\sigma)} \leq \|A\|_{L_\infty(T)}. \tag{2.3.27}$$

By (2.3.24), (2.3.26), (2.3.27) and comparability of mesh diameters (2.2.15)

$$\|J(e)\|_{L_2(\sigma)} \leq 2C_T(\bar{\gamma})^{d-1}\gamma_N^{1/2}h_T^{-1/2}\|A\|_{L_\infty(\omega_T)}\|\nabla e\|_{L_2(\omega_T)}.$$

Element T has at most $d+1$ interior true-hyperfaces yielding

$$\begin{aligned}\|J(e)\|_{L_2(\partial T)} &\leq 2(d+1)\ C_T(\bar{\gamma})^{d-1}\gamma_N^{1/2}h_T^{-1/2}\|A\|_{L_\infty(\omega_T)}\|\nabla e\|_{L_2(\omega_T)}\\ &= C_J h_T^{-1/2}\|A\|_{L_\infty(\omega_T)}\|\nabla e\|_{L_2(\omega_T)}.\end{aligned}$$

Putting together the terms from $\mathscr{L}$ and from the jump residual,

$$\begin{aligned}\eta(v,T) \leq \eta(w,T) &+ h_T\left(\|\mathrm{div}A\|_{L_\infty(T)} + C_I h_T^{-1}\|A\|_{L_\infty(T)}\right.\\ &+ \left.\|c\|_{L_\infty(T)} + \|b\|_{L_\infty(\omega)}\right)\|e\|_{H^1(T)} + h_T^{1/2} C_J h_T^{-1/2}\|A\|_{L_\infty(\omega_T)}\|e\|_{H^1(\omega_T)}\\ &\leq \eta(w,T) + C_{TOT'}\,\eta_T(D,T)\|v-w\|_{H^1(\omega_T)}\end{aligned}$$

where $C_{TOT'}$ differs by a factor of 2 for $p = 1,2$. □

2.3.4 Estimator reduction

We now establish one of the three key results we need, namely estimator reduction. This result is a minor variation of [4] Corollary 2.4 and is stated here for completeness.

Theorem 2.3.4 (Estimator reduction). *Let the problem data satisfy Assumption 2.2.1 and the mesh satisfy conditions (1) and (2) of Assumption 2.2.2. Let $\mathscr{T}_1 \in \mathbb{T}$, $\mathscr{M} \subset \mathscr{T}_1$ and $\mathscr{T}_2 = REFINE(\mathscr{T}_1, \mathscr{M})$. For $p = 1$ let*

$$\Lambda_1 := (d+2)^2\bar{\Lambda}_1^2 m_{\mathscr{E}}^{-2} \quad \textit{and} \;\; \lambda := (1-2^{-1/2d})^2 > 0$$

and for $p = 2$ let

$$\Lambda_1 := (d+2)\bar{\Lambda}_1^2 m_{\mathscr{E}}^{-2} \quad \textit{and} \;\; \lambda := 1-2^{-1/d} > 0$$

with $\bar{\Lambda}_1$ from 2.3.3 (Local Perturbation). Then for any $v_1 \in \mathbb{V}_1$ and $v_2 \in \mathbb{V}_2$ and $\delta > 0$

$$\eta_2^2(v_2,\mathscr{T}_2) \leq (1+\delta)\left\{\eta_1^2(v_1,\mathscr{T}_1) - \lambda\eta_1^2(v_1,\mathscr{M})\right\} + (1+\delta^{-1})\Lambda_1\eta_0^2|||v_2 - v_1|||^2. \tag{2.3.28}$$

Proof. The proofs for $p = 1$ and $p = 2$ are similar. For $p = 1$ it is necessary to sum over elements before squaring and for $p = 2$ square first then sum over elements.

Proof for the case $p = 1$. By the local Lipschitz property (2.3.12)

$$\eta_2(v_2,T) \leq \eta_2(v_1,T) + \bar{\Lambda}_1\eta_2(D,T)\|v_2 - v_1\|_{H^1(\omega_T)}. \tag{2.3.29}$$

Summing over all elements $T \in \mathscr{T}_2$, the sum of norms over ω_T covers each element at most $(d+2)$ times as each patch ω_T is the union of element T and the (up to) $d+1$ elements sharing a true-hyperface with T. Then by the coercivity (2.2.12) over Ω

$$\eta_2(v_2, \mathscr{T}_2) \le \eta_2(v_1, \mathscr{T}_2) + (d+2)\bar{\Lambda}_1 {m_{\mathscr{E}}}^{-1} \eta_2^2(D, \mathscr{T}_2) |||v_2 - v_1|||. \tag{2.3.30}$$

Squaring (2.3.30) and applying Young's inequality with constant δ to the cross-term,

$$\begin{aligned} \eta_2^2(v_2, \mathscr{T}_2) &\le (1+\delta)\eta_2^2(v_1, \mathscr{T}_2) + (1+\delta^{-1})(d+2)^2 \bar{\Lambda}_1^2 m_{\mathscr{E}}^{-2} \eta_2^2(D, \mathscr{T}_2) |||v_2 - v_1|||^2 \\ &= (1+\delta)\eta_2^2(v_1, \mathscr{T}_2) + (1+\delta^{-1})\Lambda_1 \eta_2^2(D, \mathscr{T}_2) |||v_2 - v_1|||^2. \end{aligned} \tag{2.3.31}$$

For an element $T \in \mathscr{M}$ marked for refinement, let $\mathscr{T}_{2,T} := \{T' \in \mathscr{T}_2 \mid T' \subset T\}$. As $v_1 \in \mathbb{V}_1$ has no discontinuities across element boundaries in $\mathscr{T}_{2,T}$, we have $J(v_1) = 0$ on true hyperfaces in the interior of $\mathscr{T}_{2,T}$.

Recall the element diameter $h_T = |T|^{1/d}$. For an element T marked for refinement, T' must be a proper subset of T, in particular a product of at least one bisection so that

$$|T'| \le \frac{1}{2}|T| \leftrightarrow |T'|^{1/d} \le \frac{1}{2^{1/d}}|T|^{1/d} \leftrightarrow h_{T'} \le \frac{1}{2^{1/d}} h_T. \tag{2.3.32}$$

Then

$$\begin{aligned} \sum_{T' \in \mathscr{T}_{2,T}} \eta_2(v_1, T') &\le \sum_{T' \in \mathscr{T}_{2,T}} h_{T'} \|R(v_1)\|_{L_2(T')} + \sum_{T' \in \mathscr{T}_{2,T}} h_{T'}^{1/2} \|J(v)\|_{L_2(\partial T' \cap \partial T)} \\ &\le 2^{-1/d} h_T \sum_{T' \in \mathscr{T}_{2,T}} \left(\|R(v_1)\|_{L_2(T')} \right) + 2^{-1/2d} h_T^{1/2} \|J(v)\|_{L_2(\partial T)} \\ &\le 2^{-1/2d} \left(h_T \|R(v_1)\|_{L_2(T)} + h_T^{1/2} \|J(v)\|_{L_2(\partial T)} \right) \\ &= 2^{-1/2d} \eta_1(v_1, T). \end{aligned} \tag{2.3.33}$$

For an element $T \notin \mathscr{M}$, that is $T' = T$ the indicator is reproduced

$$\eta_2(v_1, T') = \eta_1(v_1, T). \tag{2.3.34}$$

Sum over all $T \in \mathscr{T}_2$ by estimates (2.3.33), (2.3.34) writing the sum of indicators over the $\mathscr{T}_1 \setminus \mathscr{M}$ as the total estimator less the indicators over the refinement set $\mathscr{M}$. Let the refined set $\mathscr{R} := \{T \in \mathscr{T}_2 \bigm| T' \subset \tilde{T} \text{ for some } \tilde{T} \in \mathscr{M}\}$ then

$$\begin{aligned}
\eta_2(v_1, \mathscr{T}_2) &= \sum_{T \in \mathscr{T}_2} \eta_2(v_1, T) \\
&= \sum_{T \in \mathscr{T}_2 \setminus \mathscr{R}} \eta_2(v_1, T) + \sum_{T \in \mathscr{R}} \eta_2(v_1, T) \\
&\leq \eta_1(v_1, \mathscr{T}_1) - \eta_1(v_1, \mathscr{M}) + 2^{-1/2d} \eta_1(v_1, \mathscr{M}) \\
&= \eta_1(v_1, \mathscr{T}_1) - \lambda_1 \, \eta_1(v_1, \mathscr{M}) \qquad (2.3.35)
\end{aligned}$$

where $\lambda_1 = 1 - 2^{-1/2d} < 1$. Squaring (2.3.35)

$$\begin{aligned}
\eta_2^2(v_1, \mathscr{T}_2) &\leq \eta_1^2(v_1, \mathscr{T}_1) + \lambda_1^2 \, \eta_1^2(v_1, \mathscr{M}) - 2\lambda_1^2 \, \eta_1^2(v_1, \mathscr{M}) \\
&= \eta_1^2(v_1, \mathscr{T}_1) - \lambda \, \eta_1^2(v_1, \mathscr{M}) \qquad (2.3.36)
\end{aligned}$$

where $\lambda = \lambda_1^2 = (1 - 2^{-1/2d})^2$. Applying (2.3.36) to (2.3.31) and applying monotonicity of the data-estimator

$$\begin{aligned}
\eta_2^2(v_2, \mathscr{T}_2) \leq &(1 + \delta) \left(\eta_1^2(v_1, \mathscr{T}_1) - \lambda \, \eta_1^2(v_1, \mathscr{M}) \right) \\
&+ (1 + \delta^{-1}) \Lambda_1^2 \eta_0^2(D, \mathscr{T}_0) |||v_2 - v_1|||^2.
\end{aligned}$$

The proof for the case $p = 2$ is similar and may be found in [4]. □

2.3.5 Contraction of AFEM

We now establish the main contraction results. The contraction result 2.3.5 is a modification of [4] Theorem 4.1. Here we use quasi-orthogonality to establish contraction of each of the nonsymmetric problems (2.1.3) and (2.2.1).

Theorem 2.3.5 (GOAFEM contraction)**.** *Let the problem data satisfy Assumption 2.2.1 and the mesh satisfy Assumption 2.2.2. Let u the solution to* (2.1.3)*. Let $\theta \in (0, 1]$, and let $\{\mathscr{T}_k, \mathbb{V}_k, u_k\}_{k \geq 0}$ be the sequence of meshes, finite element spaces and discrete solutions produced by GOAFEM. Then there exist constants $\gamma > 0$ and $0 < \alpha < 1$, depending* *on the initial mesh $\mathscr{T}_0$ and marking parameter θ such that*

$$|||u-u_{k+1}|||^2+\gamma\eta_{k+1}^2 \leq \alpha^2\left(|||u-u_k|||^2+\gamma\eta_k^2\right). \tag{2.3.37}$$

The analogous result holds for the dual problem with $\{\mathscr{T}_k, \mathbb{V}_k, z_k\}_{k\geq 0}$ the sequence of meshes, finite element spaces and discrete solutions produced by GOAFEM.

Proof. Denote

$$e_k = u-u_k, \quad e_{k+1} = u-u_{k+1} \quad \text{and} \quad \varepsilon_k = u_{k+1}-u_k.$$

Let

$$\eta_k = \eta_k(u_k, \mathscr{T}_k), \quad \eta_k(\mathscr{M}_k) = \eta_k(u_k, \mathscr{M}_k) \quad \text{and} \quad \eta_{k+1} = \eta_{k+1}(u_{k+1}, \mathscr{T}_{k+1}).$$

By the result of Estimator Reduction 2.3.4, for any $\delta > 0$

$$\eta_{k+1}^2 \leq (1+\delta)\left\{\eta_k^2-\lambda\eta_k^2(\mathscr{M}_k)\right\}+(1+\delta^{-1})\Lambda_1\eta_0^2|||\varepsilon_k|||^2.$$

Multiplying this inequality by positive constant γ (to be determined) and adding the quasi-orthogonality estimate $|||e_{k+1}|||^2 \leq \Lambda|||e_k|||^2-|||\varepsilon_k|||^2$ obtain

$$\begin{aligned}|||e_{k+1}|||^2+\gamma\eta_{k+1}^2 \leq \Lambda|||e_k|||^2-|||\varepsilon_k|||^2+\gamma(1+\delta)\left\{\eta_k^2-\lambda\eta_k^2(\mathscr{M}_k)\right\}\\ +\gamma(1+\delta^{-1})\Lambda_1\eta_0^2|||\varepsilon_k|||^2.\end{aligned} \tag{2.3.38}$$

Choose γ to eliminate $|||\varepsilon_k|||$ the error between consecutive estimates by setting

$$\gamma(1+\delta^{-1})\Lambda_1\eta_0^2 = 1 \iff \gamma = \frac{1}{(1+1/\delta)\Lambda_1\eta_0^2} \iff \gamma(1+\delta) = \frac{\delta}{\Lambda_1\eta_0^2}. \tag{2.3.39}$$

Applying (2.3.39) to (2.3.38) obtain

$$|||e_{k+1}|||^2+\gamma\eta_{k+1}^2 \leq \Lambda|||e_k|||^2+\gamma(1+\delta)\eta_k^2-\gamma(1+\delta)\lambda\eta_k^2(\mathscr{M}_k). \tag{2.3.40}$$

By the Dörfler marking strategy $\eta_k^2(\mathscr{M}_k) \geq \theta^2 \eta_k^2$ so that

$$|||e_{k+1}|||^2 + \gamma \eta_{k+1}^2 \leq \Lambda |||e_k|||^2 + \gamma(1+\delta)\eta_k^2 - \gamma(1+\delta)\lambda\theta^2\eta_k^2. \tag{2.3.41}$$

Split the last term by factors of β and $(1-\beta)$ for any $\beta \in (0,1)$ to arrive at

$$\begin{aligned} |||e_{k+1}|||^2 + \gamma \eta_{k+1}^2 \leq \Lambda |||e_k|||^2 + \gamma(1+\delta)\eta_k^2 - \beta\gamma(1+\delta)\lambda\theta^2\eta_k^2 \\ -(1-\beta)\gamma(1+\delta)\lambda\theta^2\eta_k^2. \end{aligned} \tag{2.3.42}$$

Applying the upper-bound estimate (2.3.11) $|||e_k|||^2 \leq C_1 \eta_k^2$ to the term multiplied by β then by (2.3.39)

$$\begin{aligned} |||e_{k+1}|||^2 + \gamma \eta_{k+1}^2 &\leq \Lambda |||e_k|||^2 - \frac{\beta\gamma(1+\delta)\lambda\theta^2}{C_1} |||e_k|||^2 + \gamma(1+\delta)\eta_k^2 \\ &\quad -(1-\beta)\gamma(1+\delta)\lambda\theta^2\eta_k^2 && (2.3.43) \\ &= \Lambda |||e_k|||^2 - \beta \frac{\delta\lambda\theta^2}{C_1\Lambda_1\eta_0^2} |||e_k|||^2 + \gamma(1+\delta)\eta_k^2 \\ &\quad -(1-\beta)\gamma(1+\delta)\lambda\theta^2\eta_k^2 && (2.3.44) \\ &= \left(\Lambda - \beta \frac{\delta\lambda\theta^2}{C_1\Lambda_1\eta_0^2}\right) |||e_k|||^2 + \gamma(1+\delta)\left(1-(1-\beta)\lambda\theta^2\right)\eta_k^2 \\ & && (2.3.45) \\ &= \alpha_1^2(\delta,\beta) |||e_k|||^2 + \gamma\alpha_2^2(\delta,\beta)\eta_k^2 && (2.3.46) \end{aligned}$$

where

$$\alpha_1^2(\delta,\beta) := \Lambda - \beta \frac{\lambda\theta^2}{C_1\Lambda_1\eta_0^2}\delta, \quad \alpha_2^2(\delta,\beta) := (1+\delta)\left(1-(1-\beta)\lambda\theta^2\right). \tag{2.3.47}$$

Choose δ small enough so that

$$\alpha^2 := \max\{\alpha_1^2, \alpha_2^2\} < 1.$$

To ensure such a δ exists in light of the quasi-orthogonality constant $\Lambda > 1$ observe

$$\alpha_1^2 < 1 \text{ when } \delta > (\Lambda - 1)\frac{C_1\Lambda_1\eta_0^2}{\beta\lambda\theta^2}$$

and

$$\alpha_2^2 < 1 \text{ when } \delta < \left(1-(1-\beta)\lambda\theta^2\right)^{-1} - 1 = \frac{(1-\beta)\lambda\theta^2}{1-(1-\beta)\lambda\theta^2}$$

so to obtain an interval of positive measure where δ may be found we require

$$(\Lambda - 1)\frac{C_1\Lambda_1\eta_0^2}{\beta\lambda\theta^2} < \frac{(1-\beta)\lambda\theta^2}{1-(1-\beta)\lambda\theta^2}$$

placing a second constraint on the quasi-orthogonality constant

$$\Lambda < 1 + \frac{\lambda^2\theta^4\beta(1-\beta)}{C_1\Lambda_1\eta_0^2(1-(1-\beta)\lambda\theta^2)} \tag{2.3.48}$$

where $0 < \beta < 1$ and $\theta < 1$ may be chosen. In order to place bounds on the growth rate of the mesh, we further require $\theta < \theta_*$ given by (2.4.5) as discussed in section §2.4. □

Notice the choice of δ small enough to satisfy $\alpha^2 < 1$ is always possible, as each term may be independently driven below unity by a sufficiently small value of δ, so long as the quasi-orthogonality constant Λ is sufficiently close to one. For a discussion on the optimal contraction factor see Remark 4.3 in [4]; see also the discussion in [8].

2.3.6 Convergence of GOAFEM

We now derive a bound on error in the goal function.

Theorem 2.3.6 (GOAFEM functional convergence)**.** *Let the problem data satisfy Assumption 2.2.1 and the mesh satisfy Assumption 2.2.2. Let u the solution to* (2.1.3) *and z the solution to* (2.2.1). *Let $\theta \in (0,1]$, and let $\{\mathcal{T}_k, \mathbb{V}_k, u_k, z_k\}_{k\geq 0}$ be the sequence of meshes, finite element spaces and discrete primal and dual solutions produced by GOAFEM. Let γ_p the constant γ from Theorem 2.3.5 applied to the primal problem* (2.2.22) *and γ_d the constant γ from Theorem 2.3.5 applied to the dual* (2.2.23)*. Then for constant $\alpha < 1$ as determined by Theorem 2.3.5*

$$|g(u)-g(u_k)| \leq 2\left\{\alpha^{2k}\left(|||u-u_0|||^2+\gamma_p\eta_0^2(u_0,\mathscr{T}_0)\right)-\gamma_p\eta_k^2\right\}^{1/2}$$
$$\times\left\{\alpha^{2k}\left(|||z-z_0|||^2+\gamma_d\zeta_0^2(z_0,\mathscr{T}_0)\right)-\gamma_d\zeta_k^2\right\}^{1/2}.$$

Proof. On the primal side for all $v_k \in \mathbb{V}_k$

$$a(u-u_k,v_k)=a(u,v_k)-a(u_k,v_k)=f(v_k)-f(v_k)=0,$$

the primal Galerkin orthogonality property. On the dual side, $g(u)=a^*(z,u)$ and $g(u_k)=a^*(z,u_k,)$ so that

$$\begin{aligned} g(u)-g(u_k) &= a^*(z,u-u_k) \\ &= a(u-u_k,z) \\ &= a(u-u_k,z-z_k). \end{aligned} \tag{2.3.49}$$

Define an inner-product α by the symmetric part of $a(\,\cdot\,,\,\cdot\,)$

$$\alpha(v,w)=\langle A\nabla v,\nabla w\rangle+\langle cv,w\rangle,$$

then

$$|||v|||^2=a(v,v)=\alpha(v,v),$$

and

$$a(v,w)=\alpha(v,w)+\langle b\cdot\nabla v,w\rangle.$$

Then as $\alpha(\cdot,\cdot)$ is a symmetric bilinear form on Hilbert space; it is an inner product and it induces a norm identical to the energy norm induced by $a(\,\cdot\,,\,\cdot\,)$. As such we may apply the Cauchy-Schwarz inequality [6] to α and we're left to handle the convection term.

$$\begin{aligned} a(u-u_k,z-z_k) &= \alpha(u-u_k,z-z_k)+\langle b\cdot\nabla(u-u_k),z-z_k\rangle \\ &\leq |||u-u_k|||\,|||z-z_k|||+\langle b\cdot\nabla(u-u_k),z-z_k\rangle. \end{aligned} \tag{2.3.50}$$

By Hölder's inequality followed by a duality estimate as in §2.6 on the dual error and coercivity on the primal,

$$\langle b\cdot\nabla(u-u_k), z-z_k\rangle \leq \|b\|_{L_\infty} C_* h_0^s \mu_0^{-1/2} |||z-z_k||| \, |||u-u_k|||. \tag{2.3.51}$$

Recalling $\bar{\Lambda} = \|b\|_{L_\infty} C_* h_0^s \mu_0^{-1/2}$

$$a(u-u_k, z-z_k) \leq |||u-u_k||| \, |||z-z_k||| + \bar{\Lambda} |||u-u_k||| \, |||z-z_k|||. \tag{2.3.52}$$

Under assumption (2.3.6) ($\bar{\Lambda} < 1$) on the initial mesh and from (2.3.49),

$$|g(u)-g(u_k)| = |a(u-u_k, z-z_k)| \leq 2 |||u-u_k||| \, |||z-z_k|||. \tag{2.3.53}$$

From 2.3.5 there is an $\alpha < 1$ such that for the primal problem with estimator η_k

$$|||u-u_{k+1}|||^2 \leq \alpha^2 \left(|||u-u_k|||^2 + \gamma_p \eta_k^2 \right) - \gamma_p \eta_{k+1}^2 \tag{2.3.54}$$

and for the dual problem with estimator ζ_k

$$|||z-z_{k+1}|||^2 \leq \alpha^2 \left(|||z-z_k|||^2 + \gamma_d \zeta_k^2 \right) - \gamma_d \zeta_{k+1}^2. \tag{2.3.55}$$

Iterating, we have from (2.3.54) and (2.3.55)

$$|||u-u_k|||^2 + \gamma_p \eta_k^2 \leq \alpha^{2k} \left(|||u-u_0|||^2 + \gamma_p \eta_0^2 \right) \tag{2.3.56}$$

$$|||z-z_k|||^2 + \gamma_d \zeta_k^2 \leq \alpha^{2k} \left(|||z-z_0|||^2 + \gamma_d \zeta_0^2 \right). \tag{2.3.57}$$

From (2.3.53), (2.3.56) and (2.3.57) obtain the contraction of error in quantity of interest

$$\begin{aligned} |g(u)-g(u_k)| \leq 2 &\left\{ \alpha^{2k} \left(|||u-u_0|||^2 + \gamma_p \eta_0^2(u_0, \mathscr{T}_0) \right) - \gamma_p \eta_k^2 \right\}^{1/2} \\ &\times \left\{ \alpha^{2k} \left(|||z-z_0|||^2 + \gamma_d \zeta_0^2(z_0, \mathscr{T}_0) \right) - \gamma_d \zeta_k^2 \right\}^{1/2}, \end{aligned} \tag{2.3.58}$$

or more simply

$$\begin{aligned}|g(u)-g(u_k)|+\gamma_p\eta_k^2+\gamma_d\zeta_k^2 \le \alpha^{2k}\big(&|||u-u_0|||^2+\gamma_p\eta_0^2(u_0,\mathscr{T}_0)\\ &+|||z-z_0|||^2+\gamma_d\zeta_0^2(z_0,\mathscr{T}_0)\big) \qquad (2.3.59)\\ =\alpha^{2k}Q_0^2& \qquad (2.3.60)\end{aligned}$$

with Q_0 the quasi-error on the initial mesh.

□

2.4 Quasi-optimal cardinality of GOAFEM

In this section we establish the quasi-optimality of GOAFEM. The result in §2.4.5 follows from bounding the cardinality of the marked set for each of the primal and dual problems at each iteration as shown in Lemma 2.4.9. This is achieved by assuming the primal and dual solutions belong to appropriate approximation classes as discussed in §2.4.4, the optimality assumptions addressed in §2.4.2, and the supporting results below. Under the optimality assumptions, the error-indicator as an upper-bound on energy-error as shown in §2.4.1 and a bound for the oscillation term as the mesh is refined as shown in §2.4.2, a suitable reduction in global error between two consecutive iterations implies the respective refinement set satisfies the Dörfler property. We address the effect of quasi-orthogonality on the necessary reduction to achieve this result.

The estimator as global lower bound on total error in §2.4.1 is used to relate the total-error to the quasi-error in §2.4.5, connecting the contraction property for the quasi-error established in §2.3 to the quasi-optimality of the total error in §2.4.3 which shows the total error satisfies Céa's Lemma.

2.4.1 Estimator as global lower bound and localized upper bound

We start with two fairly standard results that will be needed in the complexity analysis. The *global lower bound* may be found in [9] Lemma 3.1 and a similar result in [10] Proposition 4.3 and Corollary 4.4. The *localized upper bound* is established in [4] Lemma 3.6.

Lemma 2.4.1 (Global lower bound)**.** *Let the problem data satisfy Assumption 2.2.1 and the mesh satisfy Assumption 2.2.2. Let $\mathscr{T}_1, \mathscr{T}_2 \in \mathbb{T}$ and $\mathscr{T}_2 \geq \mathscr{T}_1$ a full refinement. Let $u_k \in \mathbb{V}_k$ the solution to* (2.2.22)*, $k = 1,2$. Then there is a global constant $c_2 > 0$ such that*

$$c_2 \eta_1^2(u_1, \mathscr{T}_1) \leq |||u - u_1|||^2 + \operatorname{osc}_1^2(u_1, \mathscr{T}_1). \tag{2.4.1}$$

Lemma 2.4.2 (Localized upper bound)**.** *Let the problem data satisfy Assumption 2.2.1 and the mesh satisfy conditions (1) and (2) of Assumption 2.2.2. Let $\mathscr{T}_1, \mathscr{T}_2 \in \mathbb{T}$ with $\mathscr{T}_2 \geq \mathscr{T}_1$. Let $\mathscr{R} := \mathscr{R}_{\mathscr{T}_1 \to \mathscr{T}_2}$ the set of refined elements. Let $u_k \in \mathbb{V}_k$ the solution to* (2.2.22)*, $k = 1,2$. Then there is a global constant C_1 with*

$$|||u_2 - u_1|||^2 \leq C_1 \eta_1^2(u_1, \mathscr{R}). \tag{2.4.2}$$

2.4.2 Optimality assumptions and optimal marking

In this section we consider the assumptions on marking parameter θ and the marking strategy which allow us to characterize the growth of the adaptive mesh at each iteration with respect to the total error in 2.4.5.

We first consider oscillation on the refined mesh, following closely [4], Corollary 3.5.

Lemma 2.4.3 (Oscillation on refined mesh)**.** *Let the problem data satisfy Assumption 2.2.1 and the mesh satisfy condition (1) of Assumption 2.2.2. Let $\mathscr{T}_1, \mathscr{T}_2 \in \mathbb{T}$ with $\mathscr{T}_2 \geq \mathscr{T}_1$. Let $\Lambda_2 = \bar{\Lambda}_2^2 m_{\mathscr{E}}^{-2}$ with $\bar{\Lambda}_2$ from* (2.3.13)*. Then for all $v_1 \in \mathbb{V}_1$ and $v_2 \in \mathbb{V}_2$*

$$\operatorname{osc}_1^2(v_1, \mathscr{T}_1 \cap \mathscr{T}_2) \leq 2\operatorname{osc}_2^2(v_2, \mathscr{T}_1 \cap \mathscr{T}_2) + 2\Lambda_2 \operatorname{osc}_0^2 |||v_1 - v_2|||^2, \tag{2.4.3}$$

where $\operatorname{osc}_0^2 := \operatorname{osc}_{\mathscr{T}_0}^2(D, \mathscr{T}_0)$.

Proof. For all elements T in the intersection $T \in \mathscr{T}_1 \cap \mathscr{T}_2$

$$\operatorname{osc}_1(v_1, T) = \operatorname{osc}_2(v_1, T).$$

Applying this, and noting $v_1 \in \mathbb{V}_1 \subset \mathbb{V}_2$ and $\mathrm{osc}_j(D,T) \leq \mathrm{osc}_0(D,T),\ j=1,2$, we have from (2.3.13)

$$\mathrm{osc}_2(v_1,T) \leq \mathrm{osc}_2(v_2,T) + \bar{\Lambda}_2 \mathrm{osc}_0 \|v-w\|^2_{H^1(T)}.$$

Squaring and applying Young's inequality with $\varepsilon = 1$ yields

$$\mathrm{osc}_1^2(v_1,T) \leq 2\mathrm{osc}_2^2(v_2,T) + 2\bar{\Lambda}_2^2 \mathrm{osc}_0^2 \|v_1 - v_2\|^2_{H^1(T)}. \tag{2.4.4}$$

Summing over all $T \in \mathscr{T}_1 \cap \mathscr{T}_2$ and bounding the norm over $\mathscr{T}_1 \cap \mathscr{T}_2$ to the entire domain to apply the coercivity estimate (2.2.9)

$$\mathrm{osc}_1^2(v_1, \mathscr{T}_1 \cap \mathscr{T}_2) \leq 2\mathrm{osc}_2^2(v_2, \mathscr{T}_1 \cap \mathscr{T}_2) + 2\Lambda_2 \mathrm{osc}_0^2 |||v_1 - v_2|||^2.$$

□

We now discuss some basic assumptions for complexity analysis. The optimality assumptions follow those found in [4] with modifications in (2.4.5) to account for the non-symmetric problem, the continuity of A and the goal-oriented method.

Assumption 2.4.4 (Optimality assumptions)**.** *Assume the following conditions.*

1) *The marking parameter θ satisfies $\theta \in (0, \theta_*)$ with*

$$\theta_* = \frac{c_2}{1 + C_1(1 + \bar{\Lambda} + 2\Lambda_2 \mathrm{osc}_0^2)}, \text{ with } \mathrm{osc}_0 = \mathrm{osc}_0^2(D, \mathscr{T}_0) \tag{2.4.5}$$

   and $\bar{\Lambda}$ given by (2.3.6)*. As the data oscillation given by* (2.2.35) *is identical for the primal and dual problems and the other constants depend only on global data, θ_* may be assumed the same for both the primal an dual problems.*

2) *A marked set $\mathscr{M}_k$ of optimal cardinality (up to a factor of two) is selected (see [10]).*

3) *The distribution of refinement edges on $\mathscr{T}_0$ satsifies condition (b) of section 4 in [12].*

We now consider a basic result on optimal marking. This lemma is a variation of Lemma 5.9 in [4], modified to use quasi-orthogonality 2.3.1 rather than Galerkin orthogonality.

Lemma 2.4.5 (Optimal marking)**.** *Let the problem data satisfy Assumption 2.2.1 and the mesh satisfy Assumption 2.2.2. Let* $\mathscr{T}_1, \mathscr{T}_2 \in \mathbb{T}$*. Let* $u_k \in \mathbb{V}_k$ *the solution to* (2.2.22), $k = 1,2$*. Let the marking parameter* θ *satisfy condition (1) of Assumption 2.4.4.*

Let $\mathscr{T}_2 \geq \mathscr{T}_1$ *satisfy*

$$|||u - u_2|||^2 + \mathrm{osc}_2^2 \leq \frac{\mu}{\alpha}\left(|||u - u_1|||^2 + \mathrm{osc}_1^2\right) \tag{2.4.6}$$

which implies

$$\alpha|||u - u_2|||^2 + \mathrm{osc}_2^2 \leq \mu\left(|||u - u_1|||^2 + \mathrm{osc}_1^2\right) \tag{2.4.7}$$

for $\mu := \frac{1}{2}(1 - \frac{\theta^2}{\theta_*^2})$ *and* $\alpha = (1 + \bar{\Lambda})$*,* $\bar{\Lambda} \in (0,1)$ *given by* (2.3.6) *in the quasi-orthogonality argument and where*

$$\mathrm{osc}_1 = \mathrm{osc}_1(u_1, \mathscr{T}_1), \quad \mathrm{osc}_2 = \mathrm{osc}_2(u_2, \mathscr{T}_2), \quad \textit{and } \eta_1 = \eta_1(u_1, \mathscr{T}_1).$$

Then the set $\mathscr{R} := \mathscr{R}_{\mathscr{T}_1 \to \mathscr{T}_2}$ *satisfies the Dörfler property*

$$\eta_1(u_1, \mathscr{R}) \geq \theta\eta_1(u_1, \mathscr{T}_1).$$

Proof. (See [4] Lemma 5.9). As $0 < 2\mu < 1$, multiply inequality (2.4.1) by $1 - 2\mu$ to obtain

$$(1 - 2\mu)c_2\eta_1^2 \leq |||u - u_1|||^2 + \mathrm{osc}_1^2 - 2\mu\left(|||u - u_1|||^2 + \mathrm{osc}_1^2\right).$$

Applying (2.4.7)

$$(1 - 2\mu)c_2\eta_1^2 \leq |||u - u_1|||^2 - \alpha|||u - u_2|||^2 + \mathrm{osc}_1^2 - 2\mathrm{osc}_2^2.$$

Rearranging terms obtain

$$\mathrm{osc}_1^2 - 2\mathrm{osc}_2^2 \geq (1 - 2\mu)c_2\eta_1^2 + \alpha|||u - u_2|||^2 - |||u - u_1|||^2. \tag{2.4.8}$$

By the second quasi-orthogonality estimate (2.3.9)

$$(1+\bar{\Lambda})|||u-u_2|||^2 - |||u-u_1|||^2 \geq -(1+\bar{\Lambda})|||u_1-u_2|||^2$$

where $0 < \bar{\Lambda} < 1$. Applying (2.4.2)

$$(1+\bar{\Lambda})|||u-u_2|||^2 - |||u-u_1|||^2 \geq -(1+\bar{\Lambda})C_1\eta_1^2(u_1,\mathscr{R}). \tag{2.4.9}$$

Combining (2.4.9) with (2.4.8) obtain

$$\mathrm{osc}_1^2 - 2\mathrm{osc}_2^2 \geq (1-2\mu)c_2\eta_1^2 - (1+\bar{\Lambda})C_1\eta_1^2(u_1,R). \tag{2.4.10}$$

For refined elements $T \in \mathscr{R}$ use the dominance of the estimator over the oscillation

$$\mathrm{osc}_1^2(u_1,T) \leq \eta_1^2(u_1,T).$$

For elements $T \in \mathscr{T}_1 \cap \mathscr{T}_2$ (2.4.3) yields

$$\mathrm{osc}_1^2(u_1,\mathscr{T}_1\cap\mathscr{T}_2) - 2\mathrm{osc}_2^2(u_2,\mathscr{T}_1\cap\mathscr{T}_2) \leq 2\Lambda_2\mathrm{osc}_0^2|||u_1-u_2|||^2.$$

Then

$$\mathrm{osc}_1^2(u_1,\mathscr{T}_1) - 2\mathrm{osc}_2^2(u_2,\mathscr{T}_2) \leq \eta_1^2(u_1,\mathscr{R}) + 2\Lambda_2\mathrm{osc}_0^2|||u_1-u_2|||^2.$$

Applying (2.4.2) to the last term

$$\mathrm{osc}_1^2(u_1,\mathscr{T}_1) - 2\mathrm{osc}_2^2(u_2,\mathscr{T}_2) \leq (1+2C_1\Lambda_2\mathrm{osc}_0^2)\eta_1^2(u_1,\mathscr{R}). \tag{2.4.11}$$

Rearranging terms in (2.4.11) and applying (2.4.10)

$$\eta_1^2(u_1,\mathscr{R}) \geq \frac{(1-2\mu)c_2\eta_1^2 - (1+\bar{\Lambda})C_1\eta_1^2(u_1,R)}{(1+2C_1\Lambda_2\mathrm{osc}_0^2)}.$$

Combining like terms obtain

$$\eta_1^2(u_1,\mathscr{R}) \geq \frac{(1-2\mu)c_2}{1+C_1(1+\bar{\Lambda}+2\Lambda_2\mathrm{osc}_0^2)}\eta_1^2.$$

Applying the definitions of μ and θ_* obtain the result

$$\eta_1^2(u_1,\mathscr{R}) \geq \theta^2\eta_1^2.$$

□

Due to the use of quasi-orthogonality, the required assumption (2.4.7) is stronger than

$$|||u-u_2|||^2+\mathrm{osc}_2^2 \leq \mu\left(|||u-u_1|||^2+\mathrm{osc}_1^2\right)$$

the condition in [4] for the symmetric problem, but it is also weaker than

$$|||u-u_2|||^2+\mathrm{osc}_2^2 \leq \frac{\mu}{\alpha}\left(|||u-u_1|||^2+\mathrm{osc}_1^2\right)$$

where $\alpha = 1+\bar{\Lambda} > 1$, formally similar to the symmetric estimate. We may impose this stronger condition for ease of analysis, however in practice this says that the increase in error-reduction we require of the finer mesh needs only come from the energy-norm error, not the oscillation.

We recall a standard result on the mesh overlap, see [4] Lemma 3.7.

Lemma 2.4.6 (Overlay of meshes)**.** *Let the mesh satisfy condition (1) of Assumption 2.2.2. Let $\mathscr{T}_1,\mathscr{T}_2 \in \mathbb{T}$. Then the overlay $\mathscr{T} = \mathscr{T}_1 \oplus \mathscr{T}_2$ is conforming and satisfies*

$$\#\mathscr{T} \leq \#\mathscr{T}_1+\#\mathscr{T}_2-\#\mathscr{T}_0.$$

2.4.3 Quasi-optimality of total errror

We show the total error satisfies Céa's Lemma; e.g. see [4] Lemma 5.2. This version appropriate for the non-symmetric problem relies quasi-orthogonality 2.3.1 rather than Galerkin orthogonality.

Theorem 2.4.7 (Quasi-optimality of total error)**.** *Let the problem data satisfy Assumption 2.2.1 and the mesh satisfy Assumption 2.2.2. Let $\mathscr{T}_1 \in \mathbb{T}$. Let u the solution of* (2.1.3) *and $u_1 \in \mathbb{V}_1$ the solution of* (2.2.22)*. Then there is a constant C_D depending on the initial mesh $\mathscr{T}_0$ and the problem data D such that*

$$|||u-u_1|||^2+\operatorname{osc}_1^2(u_1,\mathscr{T}_1) \leq C_D \inf_{v\in\mathbb{V}_1}\left(|||u-v|||^2+\operatorname{osc}_1^2(v,\mathscr{T}_1)\right). \tag{2.4.12}$$

Proof. For $\varepsilon > 0$ choose $v_\varepsilon \in \mathbb{V}_1$ with

$$|||u-v_\varepsilon|||^2+\operatorname{osc}_1^2(v_\varepsilon,\mathscr{T}_1) \leq (1+\varepsilon)\inf_{v\in\mathbb{V}_1}\left(|||u-v|||^2+\operatorname{osc}_1^2(v,\mathscr{T}_1)\right).$$

By (2.4.3) with $\mathscr{T}_2 = \mathscr{T}_1$ obtain

$$\operatorname{osc}_1^2(v_1,\mathscr{T}_1) \leq 2\operatorname{osc}_1^2(v_\varepsilon,\mathscr{T}_1)+2\Lambda_2\operatorname{osc}_0^2|||u_1-v_\varepsilon|||^2. \tag{2.4.13}$$

By the same reasoning as (2.3.1) obtain

$$|||u-u_1|||^2+|||u_1-v_\varepsilon|||^2 \leq \Lambda|||u-v_\varepsilon|||^2$$

which implies

$$|||u-u_1|||^2 \leq \Lambda|||u-v_\varepsilon|||^2 \text{ and } |||u_1-v_\varepsilon|||^2 \leq \Lambda|||u-v_\varepsilon|||^2. \tag{2.4.14}$$

From (2.4.13) and (2.4.14) obtain

$$\begin{aligned}|||u-u_1|||^2+\operatorname{osc}_1^2(u_1,\mathscr{T}_1) &\leq \Lambda|||u-v_\varepsilon|||^2+2\operatorname{osc}_1^2(v_\varepsilon,\mathscr{T}_1)+2\Lambda_2\operatorname{osc}_0^2|||u_1-v_\varepsilon|||^2\\ &\leq \Lambda\left(1+2\Lambda_2\operatorname{osc}_0^2\right)|||u-v_\varepsilon|||^2+2\operatorname{osc}_1(v_\varepsilon,\mathscr{T}_1).\end{aligned}$$

Set $C_D := \max\{2,\Lambda\left(1+2\Lambda_2\operatorname{osc}_0^2\right)\}$ then

$$\begin{aligned}|||u-u_1|||^2+\operatorname{osc}_1^2(u_1,\mathscr{T}_1) &\leq C_D\left(|||u-v_\varepsilon|||^2+\operatorname{osc}_1(v_\varepsilon,\mathscr{T}_1)\right)\\ &\leq C_D(1+\varepsilon)\inf_{v\in\mathbb{V}_1}\left(|||u-v|||^2+\operatorname{osc}_1^2(v,\mathscr{T}_1)\right).\end{aligned}$$

Letting $\varepsilon \to 0$ establishes the result. □

2.4.4 Approximation classes and approximation property

For problem with solution, forcing function and data (u, f, D) and dual problem (z, g, D^*), membership in an appropriate approximation class says the solution u (respectively z) may be approximated within a given tolerance by finite element approximation while the cardinality of the mesh required to achieve the tolerance satisfies (2.4.18).

For $N > 0$ let $\mathbb{T}_N$ the set of conforming triangulations generated from the initial mesh $\mathscr{T}_0$ such that the increase in cardinality is at most N

$$\mathbb{T}_N := \{\mathscr{T} \in \mathbb{T} \mid \#\mathscr{T} - \#\mathscr{T}_0 \leq N\}.$$

For $s > 0$ define the standard approximation classes for solutions based on the energy error

$$\mathscr{A}_s := \left\{ v \in \mathbb{V} \mid \sup_{N>0} (N^s \inf_{\mathscr{T}\in\mathbb{T}_N} \inf_{v_\mathscr{T}\in\mathbb{V}_\mathscr{T}} |||v - v_\mathscr{T}||| < \infty \right\} \tag{2.4.15}$$

and for L_2 data

$$\bar{\mathscr{A}_s} := \left\{ g \in L_2(\Omega) \mid \sup_{N>0} (N^s \inf_{\mathscr{T}\in\mathbb{T}_N} \|h(g - \Pi^2_{2n-2} g)\|_{L_2(\Omega)}) < \infty \right\}. \tag{2.4.16}$$

Define a measure of approximation based on the *total error*

$$\sigma(N; v, f, D) := \inf_{\mathscr{T}\in\mathbb{T}_N} \inf_{v_\mathscr{T}\in V_\mathscr{T}} \left(|||v - v_\mathscr{T}|||^2 + \mathrm{osc}^2_\mathscr{T}(v_\mathscr{T}, \mathscr{T}) \right)^{\frac{1}{2}}$$

and denote the total error of $v_\mathscr{T} \in \mathbb{V}_\mathscr{T}$ by

$$E(v, \mathscr{T}) := \left(|||v - v_\mathscr{T}|||^2 + \mathrm{osc}^2_\mathscr{T}(v_\mathscr{T}, \mathscr{T}) \right)^{1/2}$$

and the approximation class based on the total error for $s > 0$

$$\mathbb{A}_s := \left\{ (v, f, D) \mid |v, f, D|_s := \sup_{N>0} (N^s \sigma(N; v, f, D)) < \infty \right\}. \tag{2.4.17}$$

See [4] Lemma 5.3 and Lemma 5.4 for a discussion on the relation between the classes

$\mathbb{A}_s, \mathscr{A}_s$ and $\bar{\mathscr{A}_s}$. The results in this paper are developed with respect to the class $\mathbb{A}_s$ based on the total error.

Membership of the primal and dual solutions in the approximation classes $\mathbb{A}_s$ and $\mathbb{A}_t$ is applied via the use of the two properties discussed in this section.

Lemma 2.4.8 (Approximation property)**.** *Let the mesh satisfy condition (1) of Assumption 2.2.2. Let u the solution to* (2.1.3)*. Assume $u \in \mathbb{A}_s$ and $\sigma(1;u,f,D) > 0$. Then given $\varepsilon > 0$ there is a global constant C depending only on the initial mesh $\mathscr{T}_0$ and the problem data D, a partition $\mathscr{T}_\varepsilon \in \mathscr{T}$ and a $v_\varepsilon \in \mathbb{V}_{\mathscr{T}_\varepsilon}$ such that*

$$C|u,f,D|_s \geq (\#\mathscr{T}_0 - \mathscr{T}_\varepsilon)^s \varepsilon \tag{2.4.18}$$

$$E(v_\varepsilon, \mathscr{T}_\varepsilon) \leq \varepsilon. \tag{2.4.19}$$

Proof. By (2.4.17) and property of the supremum, for any $N > 0$

$$|u,f,D|_s \geq N^s \sigma(N;u,f,D) \tag{2.4.20}$$

where $N = \#\mathscr{T} - \#\mathscr{T}_0$. Given $\varepsilon > 0$ consider all $N > 0$ such that $\sigma(N;u,f,D) \geq \varepsilon$. If there is no such N, let $N_\varepsilon = 1$. By (2.4.20)

$$|u,f,D|_s \geq \sigma_1 = \frac{\sigma_1}{\varepsilon}\varepsilon \text{ where } \sigma_1 := \sigma(1;u,f,D).$$

Applying the assumption $\sigma(1;u,f,D) > 0$

$$\frac{\varepsilon}{\sigma_1}|u,f,D|_s \geq \varepsilon$$

establishing (2.4.18) with $C = \varepsilon/\sigma_1$. Also

$$\sigma(1;u,f,D) = \inf_{\mathscr{T}\in\mathbb{T}_1} \inf_{v\in\mathbb{V}_{\mathscr{T}_1}} E(v,\mathscr{T}) < \varepsilon$$

so there is $\mathscr{T}_\varepsilon \in \mathbb{T}_1$ and $v_\varepsilon \in \mathbb{V}_{\mathscr{T}_\varepsilon}$ so that $E(v_\varepsilon, \mathscr{T}_\varepsilon) \leq \varepsilon$ establishing (2.4.19). Otherwise, there is $N > 0$ with $\sigma(N;u,f,D) \geq \varepsilon$. As the infimum over the total error goes to zero as $N \to \infty$ this holds for finitely many N so define

$$K := \max\{N > 0 \mid \sigma(N; u, f, D) \geq \varepsilon\}. \tag{2.4.21}$$

By (2.4.20) and (2.4.21)

$$|u, f, D|_s \geq K^s \sigma(K; u, f, D) \geq K^s \varepsilon. \tag{2.4.22}$$

Let $N_\varepsilon = 2K$.

$$|u, f, D|_s \geq K^s \varepsilon = 2^{-s} N_\varepsilon^s \varepsilon \implies C|u, f, D|_s \geq N_\varepsilon^s \varepsilon$$

with $C = 2^s$ establishing (2.4.18). By (2.4.21) and property of the infimum with $N_\varepsilon > K$

$$\sigma(N_\varepsilon; u, f; D) = \inf_{\mathscr{T} \in \mathbb{T}_{N_\varepsilon}} \inf_{v \in \mathbb{V}_\varepsilon} E(v, \mathscr{T}) \leq \inf_{\mathscr{T} \in \mathbb{T}_{N_K}} \inf_{v \in \mathbb{V}_\varepsilon} E(v, \mathscr{T}) < \varepsilon$$

implying $E(v_\varepsilon, \mathscr{T}_\varepsilon) \leq \varepsilon$ for some $\mathscr{T}_\varepsilon \in \mathbb{T}_\varepsilon$ and a $v_\varepsilon \in \mathbb{V}_{\mathscr{T}_\varepsilon}$ establishing (2.4.19). □

2.4.5 Cardinality of $\mathscr{M}_k$ and quasi-optimality of the mesh

The results on the cardinality of $\mathscr{M}_k$ and quasi-optimality are variations on [4] Lemma 5.10 and Theorem 5.11. Here we address the goal-oriented method discussed in 2.2.3.

Lemma 2.4.9 (Cardinality of $\mathscr{M}_k$)**.** *Let the problem data satisfy Assumption 2.2.1 and the mesh satisfy Assumption 2.2.2. Assume conditions (1) and (2) of Assumption 2.4.4. Let u the solution of* (2.1.3) *and z the solution of* (2.2.1)*. Let $\{\mathscr{T}_k, \mathbb{V}_k, u_k, z_k\}_{k \geq 0}$ the sequence of meshes, finite element spaces and discrete primal and dual solutions produced by GOAFEM. If $(u, f, D) \in \mathbb{A}_s$ and $(z, g, D^*) \in \mathbb{A}_t$ we have*

$$\begin{aligned} \#\mathscr{M}_k \leq 2C \Bigg\{ & (1 + \bar{\Lambda})^{1/2s} \left(1 - \frac{\theta^2}{\theta_*^2}\right)^{-1/2s} |u, f, D|_s^{1/s} C_D^{1/2s} E_k^{-1/s}(u_k, \mathscr{T}_k) \\ & + (1 + \bar{\Lambda})^{1/2t} \left(1 - \frac{\theta^2}{\theta_*^2}\right)^{-1/2t} |z, g, D^*|_t^{1/t} C_D^{1/2t} E_k^{-1/t}(z_k, \mathscr{T}_k) \Bigg\} \end{aligned} \tag{2.4.23}$$

where C_D is the constant from (2.4.12) *and the total errors in the primal and dual prob-*

lems

$$E_k^2(u_k,\mathscr{T}_k) := |||u-u_k|||^2 + \operatorname{osc}_k^2(u_k,\mathscr{T}_k)$$
$$E_k^2(z_k,\mathscr{T}_k) := |||z-z_k|||^2 + \operatorname{osc}_k^2(z_k,\mathscr{T}_k).$$

Proof. Set $\tilde{\mu} = \frac{1}{2}\left(1-\frac{\theta^2}{\theta_*^2}\right)(1+\bar{\Lambda})^{-1}$ with $\bar{\Lambda}$ given by (2.3.6).

$$\varepsilon_p^2 := \tilde{\mu} C_D^{-1} E_k^2(u_k,\mathscr{T}_k), \text{ and } \varepsilon_d^2 := \tilde{\mu} C_D^{-1} E_k^2(z_k,\mathscr{T}_k).$$

As $(u,f,D) \in \mathbb{A}_s$, by the properties in section 2.4.4 there is a $\mathscr{T}_p \in \mathbb{T}$ and a $v_p \in \mathbb{V}_{\mathscr{T}_p}$ such that

$$\#\mathscr{T}_p - \#\mathscr{T}_0 \leq C|u,f,D|_s^{1/s}\varepsilon_p^{-1/s} \tag{2.4.24}$$
$$|||u-v_p|||^2 + \operatorname{osc}^2_{\mathscr{T}_p}(v_p,\mathscr{T}_p) \leq \varepsilon_p^2. \tag{2.4.25}$$

Similarly for $(z,g,D^*) \in \mathbb{A}_t$, there is a $\mathscr{T}_d \in \mathbb{T}$ and a $w_d \in \mathbb{V}_{\mathscr{T}_d}$ such that

$$\#\mathscr{T}_d - \#\mathscr{T}_0 \leq C|z,g,D^*|_t^{1/t}\varepsilon_d^{-1/t} \tag{2.4.26}$$
$$|||z-w_d|||^2 + \operatorname{osc}^2_{\mathscr{T}_d}(w_d,\mathscr{T}_d) \leq \varepsilon_d^2. \tag{2.4.27}$$

Let $\mathscr{T}_2 := \mathscr{T}_k \oplus (\mathscr{T}_p \oplus \mathscr{T}_d)$ as in Lemma 2.4.6. Let $u_2 \in \mathbb{V}_2$ the Galerkin solution to (2.2.22) and $z_2 \in \mathbb{V}_2$ the respective solution to (2.2.23) . See there is a reduction in the total error by a factor of $\tilde{\mu}$ from u_k to u_2 (respectively z_k to z_2). Since $\mathscr{T}_2 \geq \mathscr{T}_p$ by Theorem 2.4.7, monotonicity of infimum over total error and (2.4.25)

$$\begin{aligned} |||u-u_2|||^2 + \operatorname{osc}_2^2(u_2,\mathscr{T}_2) &\leq C_D \inf_{v\in\mathbb{V}_2}\left(|||u-v|||^2 + \operatorname{osc}_2^2(v,\mathscr{T}_2)\right) \\ &\leq C_D\varepsilon_p^2 \\ &= \tilde{\mu}\left(|||u-u_k|||^2 + \operatorname{osc}_k^2(u_k,\mathscr{T}_k)\right). \end{aligned} \tag{2.4.28}$$

Similarly for the dual problem

$$|||z-z_2|||^2 + \operatorname{osc}_2^2(z_2,\mathscr{T}_2) \leq \tilde{\mu}\left(|||z-z_k|||^2 + \operatorname{osc}_k^2(z_k,\mathscr{T}_k)\right). \tag{2.4.29}$$

This satisfies the hypothesis (2.4.6) in each problem so applying 2.4.2 the refining subset $\mathscr{R} := \mathscr{R}_{\mathscr{T}_k \to \mathscr{T}_2} \subset \mathscr{T}_k$ satisfies the Dörfler property for $\theta \leq \theta_*$. The marking procedure selects a subset for marking $\mathscr{M}_k \subset \mathscr{T}_k$ of minimal cardinality up to a factor of two so that by Lemma 2.4.6

$$\#\mathscr{M}_k \leq 2\#\mathscr{R} \leq 2(\#\mathscr{T}_2 - \#\mathscr{T}_k) \leq 2\left\{(\#\mathscr{T}_p - \#\mathscr{T}_0) + (\#\mathscr{T}_d - \#\mathscr{T}_0)\right\}. \tag{2.4.30}$$

By (2.4.30), (2.4.24), the definition of ε_p and ε_d, (2.4.28) and the definition of μ

$$\begin{aligned}
\#\mathscr{M}_k &\leq 2\left\{(\#\mathscr{T}_p - \#\mathscr{T}_0) + (\#\mathscr{T}_d - \#\mathscr{T}_0)\right\} \\
&\leq 2C\left\{|u,f,D|_s^{1/s}\varepsilon_p^{-1/s} + |z,g,D^*|_t^{1/t}\varepsilon_d^{-1/t}\right\} \\
&= 2C\left\{(1+\bar{\Lambda})^{1/2s}\left(1-\frac{\theta^2}{\theta_*^2}\right)^{-1/2s}|u,f,D|_s^{1/s}C_D^{1/2s}E_k^{-1/s}(u_k,\mathscr{T}_k)\right. \\
&\qquad \left. +(1+\bar{\Lambda})^{1/2t}\left(1-\frac{\theta^2}{\theta_*^2}\right)^{-1/2t}|z,g,D^*|_t^{1/t}C_D^{1/2t}E_k^{-1/t}(z_k,\mathscr{T}_k).\right\}
\end{aligned}$$

□

Theorem 2.4.10 (Quasi-optimality)**.** *Let the problem data satisfy Assumption 2.2.1 and the mesh satisfy Assumption 2.2.2. Let Assumption 2.4.4 be satisfied by GOAFEM. Let u the solution of* (2.1.3) *and z the solution of* (2.2.1). *Let* $\{\mathscr{T}_k, \mathbb{V}_k, u_k, z_k\}_{k\geq 0}$ *the sequence of meshes, finite element spaces and discrete primal and dual solutions produced by GOAFEM. Let* $(u,f,D) \in \mathbb{A}_s$ *and* $(z,g,D^*) \in \mathbb{A}_t$. *Then*

$$\begin{aligned}
\#\mathscr{T}_k - \#\mathscr{T}_0 \leq S_\theta &\left\{M_p\left(1+\frac{\gamma_p}{c_2}\right)^{1/2s}Q_k^{-1/s}(u_k,\mathscr{T}_k)\right. \\
&\left. + M_d\left(1+\frac{\gamma_d}{c_2}\right)^{1/2t}Q_k^{-1/t}(z_k,\mathscr{T}_k)\right\}.
\end{aligned}$$

Proof. Let the total error in primal and dual problems $E_k(u_k,\mathscr{T}_k)$ and $E_k(z_k,\mathscr{T}_k)$ as in

Lemma 2.4.9. Denote the quasi-error in each problem by

$$\begin{aligned} Q_k^2(u_k,\mathscr{T}_k) &:= |||u-u_k|||^2+\gamma_p\eta_k^2(u_k,\mathscr{T}_k),\\ Q_k^2(z_k,\mathscr{T}_k) &:= |||z-z_k|||^2+\gamma_d\zeta_k^2(z_k,\mathscr{T}_k). \end{aligned}$$

As shown in [2] Theorem 2.4 there is a global constant C_f which satisfies

$$\#\mathscr{T}_k-\#\mathscr{T}_0\le C_f\sum_{j=0}^{k-1}\#\mathscr{M}_j \quad \text{for all } k\ge 1$$

and by (2.4.23)

$$\begin{aligned} \#\mathscr{M}_k \le 2C\Bigg\{&(1+\bar{\Lambda})^{1/2s}\left(1-\frac{\theta^2}{\theta_*^2}\right)^{-1/2s}|u,f,D|_s^{1/s}C_D^{1/2s}E_k^{-1/s}(u_k,\mathscr{T}_k)\\ &+(1+\bar{\Lambda})^{1/2t}\left(1-\frac{\theta^2}{\theta_*^2}\right)^{-1/2t}|z,g,D^*|_t^{1/t}C_D^{1/2t}E_k^{-1/t}(z_k,\mathscr{T}_k)\Bigg\} \end{aligned}$$

then we have

$$\#\mathscr{T}_k-\#\mathscr{T}_0\le M_p\sum_{j=0}^{k-1}E_k(u_k,\mathscr{T}_k)^{-1/s}+M_d\sum_{j=0}^{k-1}E_k(z_k,\mathscr{T}_k)^{-1/t} \tag{2.4.31}$$

with the constants

$$\begin{aligned} M_p &:= 2C_fC(1+\bar{\Lambda})^{1/2s}\left(1-\frac{\theta^2}{\theta_*^2}\right)^{-1/2s}|u,f,D|_s^{1/s}C_D^{1/2s}\\ M_d &:= 2C_fC(1+\bar{\Lambda})^{1/2t}\left(1-\frac{\theta^2}{\theta_*^2}\right)^{-1/2t}|z,g,D^*|_t^{1/t}C_D^{1/2t}. \end{aligned}$$

From the domination of the error estimator over the oscillation and the lower bound on total error (2.4.1) we have the equivalence of the total error and quasi-error

$$\begin{aligned} |||u-u_j|||^2+\gamma_p\mathrm{osc}_j^2(u_j,\mathscr{T}_j) &\le |||u-u_j|||^2+\gamma_p\eta_j^2(u_j,\mathscr{T}_j)\\ &\le \left(1+\frac{\gamma_p}{c_2}\right)E_j^2(u_j,\mathscr{T}_j). \end{aligned} \tag{2.4.32}$$

or

$$E_j^{-1/s}(u_j,\mathscr{T}_j) \le \left(1+\frac{\gamma_p}{c_2}\right)^{1/2s} Q_j^{-1/s}(u_j,\mathscr{T}_j) \tag{2.4.33}$$

and similarly for the dual problem

$$E_j^{-1/t}(z_j,\mathscr{T}_j) \le \left(1+\frac{\gamma_d}{c_2}\right)^{1/2t} Q_j^{-1/t}(z_j,\mathscr{T}_j). \tag{2.4.34}$$

By the contraction result on the quasi-error (3.4.39) for $0 \le j \le k-1$

$$Q_k^2(u_k,\mathscr{T}_k) \le \alpha^{2(k-j)} Q_j^2(u_j,\mathscr{T}_j) \quad \text{and} \quad Q_k^2(z_k,\mathscr{T}_k) \le \alpha^{2(k-j)} Q_j^2(z_j,\mathscr{T}_j). \tag{2.4.35}$$

Putting together (2.4.31), (2.4.33) and (2.4.35) obtain

$$\begin{aligned}
\#\mathscr{T}_k - \#\mathscr{T}_0 &\le M_p \sum_{j=0}^{k-1} E_k(u_k,\mathscr{T}_k)^{-1/s} + M_d \sum_{j=0}^{k-1} E_k(z_k,\mathscr{T}_k)^{-1/t} \\
&\le \left\{ M_p \left(1+\frac{\gamma_p}{c_2}\right)^{1/2s} Q_k(u_k,\mathscr{T}_k)^{-1/s} \right. \\
&\qquad \left. + M_d \left(1+\frac{\gamma_d}{c_2}\right)^{1/2t} Q_k(z_k,\mathscr{T}_k)^{-1/t} \right\} \sum_{j=1}^{k} \alpha^{j/s}
\end{aligned}$$

where the geometric series in $\alpha < 1$ is bounded by $S_\theta = \alpha^{1/s}(1-\alpha^{1/s})^{-1}$. Then

$$\begin{aligned}
\#\mathscr{T}_k - \#\mathscr{T}_0 &\le S_\theta \left\{ M_p \left(1+\frac{\gamma_p}{c_2}\right)^{1/2s} Q_k(u_k,\mathscr{T}_k)^{-1/s} \right. \\
&\qquad \left. + M_d \left(1+\frac{\gamma_d}{c_2}\right)^{1/2t} Q_k(z_k,\mathscr{T}_k)^{-1/t} \right\} \\
&\le S_\theta \left\{ M_p \left(1+\frac{\gamma_p}{c_2}\right)^{1/2s} \left(|||u-u_k|||^2 + \gamma_p \mathrm{osc}_k^2(u_k,\mathscr{T}_k)\right)^{-1/2s} \right. \\
&\qquad \left. + M_d \left(1+\frac{\gamma_d}{c_2}\right)^{1/2t} \left(|||z-z_k|||^2 + \gamma_d \mathrm{osc}_k^2(z_k,\mathscr{T}_k)\right)^{-1/2t} \right\}.
\end{aligned}$$

As seen in (2.4.32) the total error and quasi-error are equivalent up to a constant so this result may be viewed with respect to either the quasi- or total-error. □

2.5 Conclusion

In this article we developed convergence theory for a class of goal-oriented adaptive finite element methods for second order nonsymmetric linear elliptic equations. In particular, we established contraction and quasi-optimality results for a method of this type for the elliptic problem (2.1.1)–(2.1.2) with A Lipschitz, almost-everywhere symmetric positive definite (SPD), with b divergence-free, and with $c \geq 0$. We first described the problem class in some detail, with a brief review of conforming finite element discretization and error-estimate-driven adaptive finite element methods (AFEM). We then described a goal-oriented variation of standard AFEM (GOAFEM). Following the recent work of Mommer and Stevenson [10] for symmetric problems, we established contraction of GOAFEM. We also showed convergence in the sense of the goal function. Our analysis approach was signficantly different from that of Mommer and Stevenson [10], and involved the combination of the recent contraction frameworks of Cascon et. al [4], Nochetto, Siebert, and Veeser [11], and of Holst, Tsogtgerel, and Zhu [8]. We also did a careful complexity analysis, and established quasi-optimal cardinality of GOAFEM.

Problems that were not yet addressed include allowing for jump discontinuities in the diffusion cofficient, and allowing for lower-order nonlinear terms. We will address both of these aspects in a future work.

2.6 Appendix

Duality

We include an appendix discussion of the duality argument used in the quasi-orthogonality estimate in an effort to make the paper more self-contained.

Let $u \in H_0^1(\Omega)$ the variational solution to (2.1.3) and $u_1 \in \mathbb{V}_1$ the Galerkin solution to (2.2.22). Assume for any $g \in L_2(\Omega)$ the solution w to the dual problem (2.2.1) belongs to $H^2(\Omega) \cap H_0^1(\Omega)$ and

$$|w|_{H^2(\Omega)} \leq K_R \|g\|_{L_2(\Omega)}. \tag{2.6.1}$$

Then

$$\|u-u_1\|_{L_2} \leq Ch_0 |\!|\!| u-u_1 |\!|\!|. \tag{2.6.2}$$

If $w \in H^2_{\text{loc}}(\Omega) \cap H^1_0(\Omega)$ but $w \notin H^2(\Omega)$ due to the angles of a nonconvex polyhedral domain Ω then $w \in H^{1+s}$ for some $0 < s < 1$ where s depends on the angles of $\partial\Omega$. Assume in this case for any $g \in L_2$

$$|w|_{H^{1+s}(\Omega)} \leq K_R \|g\|_{L_2(\Omega)} \tag{2.6.3}$$

then

$$\|u-u_1\|_{L_2} \leq Ch_0^s |\!|\!| u-u_1 |\!|\!|. \tag{2.6.4}$$

As discussed in [5], [6] and [1] the regularity assumptions are reasonable based on the continuity of the diffusion coefficients a_{ij} and the convection and reaction coefficients b_i and c in $L_\infty(\Omega)$.

Proof of (2.6.2)*:* The proof follows the duality arguments in [1] and [3].

Let $w \in H^1_0(\Omega)$ the solution to the dual problem

$$a^*(w,v) = \langle u-u_1, v\rangle, \quad v \in H^1_0(\Omega). \tag{2.6.5}$$

Let $\mathscr{I}^h$ a global interpolator based on refinement $\mathscr{T}_1$. Assume $\mathscr{I}^h w$ is C^0 and the corresponding shape functions have approximation order m. For $m = 2$

$$\|w - \mathscr{I}^h w\|_{H^1} \leq C_{\mathscr{I}} h_{\mathscr{T}_1} |w|_{H^2.} \tag{2.6.6}$$

As discussed in [1] the interpolation estimate over reference element $\hat{T}$ follows from the Bramble-Hilbert lemma applied to the bounded linear functional $f(\hat{u}) = \langle \hat{u} - \mathscr{I}^h \hat{u}, \hat{v}\rangle$ where $\hat{v} \in H^t(\hat{T})$ is arbitrary then set to $\hat{u} - \mathscr{I}^h \hat{u}$. The Sobolev semi-norms for $t = 0,1$ over elements $T \in \mathscr{T}$ are bounded via change of variables to the reference element. Summing over $T \in \mathscr{T}$ and combining semi-norms into a norm estimate establishes (2.6.6).

By (2.6.1) we have the bound

$$|w|_{H^2} \leq K_R \|u-u_1\|_{L_2}. \tag{2.6.7}$$

By the identity $a(v,y) = a^*(y,v)$ write the primal form of the variational problems

$$a(u,v) = f(v), \quad v \in H_0^1(\Omega) \tag{2.6.8}$$

$$a(u_1,v) = f(v), \quad v \in \mathbb{V}_1 \tag{2.6.9}$$

$$a(v,w) = \langle u - u_1, v \rangle, \quad v \in H_0^1(\Omega). \tag{2.6.10}$$

Taking $v = u - u_1 \in H_0^1$ in (2.6.10)

$$a(u - u_1, w) = \langle u - u_1, u - u_1 \rangle = \|u - u_1\|_{L_2}^2. \tag{2.6.11}$$

Combining (2.6.8) and (2.6.9) we have the Galerkin orthogonality result

$$a(u - u_1, v) = 0, \quad v \in \mathbb{V}_1. \tag{2.6.12}$$

Then by (2.6.11) and (2.6.12) noting the interpolant of the dual solution $\mathscr{I}^h w \in \mathbb{V}_1$

$$\|u - u_1\|_{L_2}^2 = a(u - u_1, w) = a(u - u_1, w - \mathscr{I}^h w). \tag{2.6.13}$$

Starting with (2.6.13) and applying continuity (2.2.8), interpolation estimate (2.6.6) and elliptic regularity (2.6.7)

$$\begin{aligned}
\|u - u_1\|_{L_2}^2 &\leq M_c \|u - u_1\|_{H^1} \|w - \mathscr{I}^h w\|_{H^1} \\
&\leq M_c \|u - u_1\|_{H^1} C_{\mathscr{I}} h_{\mathscr{T}_1} |w|_{H^2} \\
&\leq K_R M_c C_{\mathscr{I}} h_0 \|u - u_1\|_{H^1} \|u - u_1\|_{L_2}.
\end{aligned}$$

Canceling one factor of $\|u - u_1\|_{L_2}$ and applying coercivity (2.2.9)

$$\|u - u_1\|_{L_2} \leq \frac{M_c}{m_{\mathscr{E}}} C_{\mathscr{I}} K_R h_0 |||u - u_1|||. \tag{2.6.14}$$

Depending on the regularity of the boundary $\partial\Omega$ the solution w may have less regularity: $w \in H^2_{\mathrm{loc}(\Omega)}$ but $w \notin H^2(\Omega)$. In particular, we may have $w \in H^{1+s}$ for some

$s \in (0,1)$. In that case obtain the more general estimate

$$\|w - \mathscr{I}^h w\|_{H^1} \leq \tilde{C}_{\mathscr{I}} h_0^s |w|_{1+s}$$

yielding

$$\|u - u_1\|_{L_2} \leq \frac{M_c}{m_{\mathscr{E}}} \tilde{C}_{\mathscr{I}} K_R h_0^s |||u - u_1|||.$$

The value of s is found by considering all corners of boundary $\partial\Omega$. Writing the interior angle at each corner by $\omega = \pi/\alpha$ it holds for $\alpha > 0$ and arbitrary $\varepsilon > 0$

$$\omega = \pi/\alpha \implies w \in H^{1+\alpha-\varepsilon}$$

and if $\pi/(p_j+1) \leq \omega \leq \pi/p_j$ for a set of integers p_j characterizing the corners of $\partial\Omega$

$$\|w - \mathscr{I}^h w\|_{H^1} \leq Ch^s |w|_{1+s}$$

where $s = \min\{p_j, 1\}$ and $s = 1$ in the case of a smooth boundary or a convex polyhedral domain. Details may be found in [1] and [13].

Acknowledgments

MH was supported in part by NSF Awards 0715146 and 0915220 and by DOD/DTRA Award HDTRA-09-1-0036. SP was supported in part by NSF Award 0715146.

References

[1] O. Axelsson and V. A. Barker. *Finite element solution of boundary value problems: theory and computation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.

[2] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.

[3] S. Brenner and L. Scott. *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, third edition, 2008.

[4] J. M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method, in preparation. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.

[5] P. G. Ciarlet. *Finite Element Method for Elliptic Problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.

[6] L. C. Evans. *Partial Differential Equations (Graduate Studies in Mathematics, V. 19) GSM/19*. American Mathematical Society, 1998.

[7] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Springer-Verlag, 1977.

[8] M. Holst, G. Tsogtgerel, and Y. Zhu. Local and global convergence of adaptive methods for nonlinear partial differential equations, 2008.

[9] K. Mekchay and R. Nochetto. Convergence of adaptive finite element methods for general second order linear elliptic PDE. *SINUM*, 43(5):1803–1827, 2005.

[10] M. S. Mommer and R. Stevenson. A goal-oriented adaptive finite element method with convergence rates. *SIAM J. Numer. Anal.*, 47(2):861–886, 2009.

[11] R. H. Nochetto, K. G. Siebert, and A. Veeser. *Theory of adaptive finite element methods: an introduction*, pages 409 – 542. Springer, 2009.

[12] R. Stevenson. The completion of locally refined simplicial partitions created by bisection. Technical report, 2006.

[13] G. Strang and G. J. Fix. *An Analysis of the Finite Element Method*. Prentice-Hall (Series in Automatic Computation), Englewood Cliffs, N. J., 1973.

Chapter 2, in full, has been submitted for publication of the material as it may appear in Applied Journal of Numerical Mathematics, 2011, M. Holst and S. Pollock, Elsevier, 2011. The dissertation author was the primary investigator and author of this paper.

Chapter 3

Convergence of Goal-Oriented Adaptive Finite Element Methods for Semilinear Problems

Convergence of Goal-Oriented Adaptive Finite Element Methods for Semilinear Problems

Michael Holst, Sara Pollock, and Yunrong Zhu

Abstract. In this article we develop convergence theory for a class of goal-oriented adaptive finite element algorithms for second order semilinear elliptic equations. We first introduce several approximate dual problems, and briefly discuss the target problem class. We then review some standard facts concerning conforming finite element discretization and error-estimate-driven adaptive finite element methods (AFEM). We include a brief summary of *a priori* estimates for semilinear problems, and then describe goal-oriented variations of the standard approach to AFEM (GOAFEM). Following the recent approach of Mommer-Stevenson and Holst-Pollock for linear problems, we establish a contraction result for the primal problem. We then develop some additional estimates that make it possible to establish contraction of the combined primal-dual quasi-error, and subsequently show convergence in the sense of the quantity of interest. Numerical experiments support the theoretical results.

3.1 Introduction

In this article we develop convergence theory for a class of goal-oriented adaptive finite element methods for second order semilinear equations. In particular, we establish strong contraction results for a method of this type for the problem:

$$\begin{cases} \mathcal{N}(u) := -\nabla \cdot (A\nabla u) + b(u) = f, & \text{in } \Omega, \\ \hfill u = 0, & \text{on } \partial\Omega, \end{cases} \tag{3.1.1}$$

with $f \in L_2(\Omega)$ and $\Omega \subset \mathbb{R}^d$ ($d = 2$ or 3) a polyhedral domain. We consider the problem with $A\colon \Omega \to \mathbb{R}^{d\times d}$ Lipschitz and almost-everywhere (a.e.) symmetric positive definite (SPD). The standard weak formulation of the primal problem reads: Find $u \in H_0^1(\Omega)$ such that

$$\langle \mathcal{N}(u), v\rangle := a(u,v) + \langle b(u), v\rangle = f(v), \quad \forall v \in H_0^1(\Omega), \tag{3.1.2}$$

where

$$a(u,v) = \int_\Omega A\nabla u \cdot \nabla v \, dx. \tag{3.1.3}$$

In many practical applications, one is more interested in certain physical quantities of the solution, referred to as “quantities of interest”, such as (weighted) averages, flow rates or velocities. These quantities of interest are usually charactered by the value $g(u)$, where u is the solution and g is a linear functional associated with a particular “goal”. Given a numerical approximation u_h to the solution u, goal-oriented error estimates use duality techniques rather than the energy norm alone to estimate the error in the quantity of interest . The solution of the dual problem can be interpreted as the generalized Green’s function, or the *influence function* with respect to the linear functional, which often quantifies the stability properties of the computed solution. There has been a lot of recent work on developing reliable and accurate *a posteriori* error estimators for goal-oriented adaptivity, see [12, 4, 5, 14, 35, 13, 18, 19, 30] and the references cited therein.

Our interest in this paper is developing a goal-oriented adaptive algorithm for semilinear problems (3.1.2) along with a corresponding strong contraction result, following the recent approach in [32, 23] for linear problems. One of the main challenges in the nonlinear problem that we don't see in the linear case is the dependence of the dual problem on the primal solution u. As it is not practical to work with a dual problem we cannot accurately form, we develop a method for semilinear problems in which adaptive mesh refinement is driven both by residual-based approximation to the error in u, and in a sequence of *approximate dual problems* which only depend on the numerical solution from the previous step. While globally reducing the error in the primal problem necessarily yields a good approximation to the goal error $g(u-u_h)$, methods of the type we describe here bias the error reduction in the direction of the goal-function g in the interest of achieving an accurate approximation in fewer adaptive iterations.

Contraction of the adaptive finite element algorithm for the (primal) semilinear problem (3.1.2) has been established in [25] and [22]. Here we recall the contraction argument for the primal problem and use a generalization of this technique to establish the contraction of a linear combination of the primal and limiting dual quasi-errors by means of a computable sequence of approximate dual problems. We relate this result to a bound on the error in the quantity of interest. Following [25], the contraction argument follows from first establishing three preliminary results for two successive AFEM approximations u_1 and u_2, and respectively $\hat{z}_1$ and $\hat{z}_2$ of the primal and limiting dual problems.

1) Quasi-orthogonality: There exists $\Lambda_G > 1$ such that

$$|||u-u_2|||^2 \leq \Lambda_G |||u-u_1|||^2 - |||u_2-u_1|||^2.$$

2) Error estimator as upper bound on error: There exists $C_1 > 0$ such that

$$|||u-u_k|||^2 \leq C_1 \eta_k^2(u_k, \mathscr{T}_k), \quad k = 1, 2.$$

3) Estimator reduction: For $\mathscr{M}$ the marked set that takes refinement $\mathscr{T}_1 \to \mathscr{T}_2$, for positive constants $\lambda < 1$ and Λ_1 and any $\delta > 0$

$$\eta_2^2(v_2, \mathscr{T}_2) \leq (1+\delta)\{\eta_1^2(v_1, \mathscr{T}_1) - \lambda \eta_1^2(v_1, \mathscr{M})\} + (1+\delta^{-1})\Lambda_1 \eta_0^2 |||v_2 - v_1|||.$$

For the primal problem, the mesh at each iteration may be marked for refinement with respect to the error indicators following the Dörfler marking strategy (cf. [11]). In the case of the dual, the limiting estimator as used in the contraction argument is related to a computable quantity. This quantity is the dual estimator, based on the residual of the approximate dual sequence. The mesh is marked for refinement with respect to this set of error indicators, which correspond to the approximate dual problem at each iteration. The transformation between limiting and approximate dual estimators couples the contraction of error in the limiting dual to the primal problem. The final result is the contraction of what we refer to here as the *combined quasi-error*

$$\bar{Q}^2(u_j,\hat{z}_j) := |||\hat{z}-\hat{z}_j|||^2 + \gamma\zeta_2^2(\hat{z}_j) + \pi|||u-u_j|||^2 + \pi\gamma_p\eta_2^2(u_j),$$

which is the sum of the quasi-error as in [8] for the limiting dual problem and a multiple of the quasi-error for the primal problem. The contraction of this property as shown in Theorem 3.5.9 establishes the contraction of the error in the goal function as shown in Corollary 3.5.10.

Our analysis is based on the recent contraction framework for semilinear and more general nonlinear problems developed by Holst, Tsogtgerel, and Zhu [25], and by Bank, Holst, Szypowski and Zhu [2], and those for linear problems developed by Cascon, Kreuzer, Nochetto and Siebert [8], and by Nochetto, Siebert, and Veeser [34]. In addressing the goal-oriented problem we base our framework on that of Mommer and Stevenson [32] for symmetric linear problems and Holst and Pollock [23] for non-symmetric problems. We note also two other recent convergence results in the literature for goal-oriented adaptive methods applied to self-adjoint linear problems, namely [10] and [33], both providing convergence rates in agreement with those in [32]. The analysis of the goal-oriented method for nonlinear problems is signficantly more complex than the previous analysis for linear problems in [32, 23]. Here, we are faced with analyzing linearized and approximate dual sequences as opposed to a single dual problem in order to establish contraction with respect to the quantity of interest. However, this approach allows us to establish a contraction result for the goal-oriented method, which appears to be the first result of this type for nonlinear problems. The linearized dual in the context of goal-oriented adaptive methods is described below, following e.g. Estep et. al in [14] and [15].

Outline of the paper. The remainder of the paper is structured as follows. In §3.2, we introduce the approximate, linearized and limiting dual problems. We briefly discuss the problem class and review some standard facts concerning conforming finite element discretization and error-estimate-driven adaptive finite element methods (AFEM). In §3.2.2 we include a brief summary of *a priori* estimates for the semilinear problem. In §3.3, we describe a goal-oriented variation of the standard approach to AFEM (GOAFEM). In §3.4 we discuss contraction theorems for the primal problem. In §3.5 we introduce additional estimates necessary for the contraction of the combined quasi-error and convergence in the sense of the quantity of interest. Lastly, in §3.6 we present some numerical experiments that support our theoretical results.

3.2 Preliminaries

In this section, we state both the (nonlinear) primal problem and its finite element discretization. We then introduce the linearized dual problem, and consider some variants of this problem which are of use in the subsequent computation and analysis.

Consider the semilinear problem (3.1.2), where as in (3.1.3) we define the bilinear form

$$a(u,v) = (A\nabla u, \nabla v)$$

with $(\cdot,\cdot)$ denoting the L_2 inner-product over $\Omega \subset \mathbb{R}^d$. We make the following assumptions on the data:

Assumption 3.2.1 (Problem data)**.** *The problem data* $\mathbf{D} = (A, b, f)$ *satisfies*

1) $A : \Omega \to \mathbb{R}^{d\times d}$ *is Lipschitz continuous and a.e. symmetric positive-definite with*

$$\begin{aligned} ess\ inf_{x\in\Omega}\lambda_{min}(A(x)) &= \mu_0 > 0, \\ ess\ sup_{x\in\Omega}\lambda_{max}(A(x)) &= \mu_1 < \infty. \end{aligned}$$

2) $b : \Omega \times \mathbb{R} \to \mathbb{R}$ *is smooth on the second argument. Here and in the remainder of the paper, we write* $b(u)$ *instead of* $b(x,u)$ *for simplicity. Moreover, we assume* *that b is monotone (increasing):*

$$b'(\xi) \geq 0, \ \textit{for all } \xi \in \mathbb{R}.$$

3) $f \in L_2(\Omega)$.

The native norm is the Sobolev H^1 norm given by $\|v\|_{H^1}^2 = (\nabla v, \nabla v) + (v,v)$. Continuity of $a(\,\cdot\,,\,\cdot\,)$ follows from the Hölder inequality, and bounding the L_2 norm of the function and its gradient by the H^1 norm

$$a(u,v) \leq \mu_1 \|u\|_{H^1} \|v\|_{H^1} = M_{\mathscr{E}} \|u\|_{H^1} \|v\|_{H^1}. \tag{3.2.1}$$

Define the energy semi-norm by the principal part of the differential operator $|||v|||^2 := a(v,v)$. The coercivity of $a(\,\cdot\,,\,\cdot\,)$ follows from the Poincaré inequality with constant C_Ω

$$a(v,v) \geq \mu_0 |v|_{H^1}^2 \geq C_\Omega \mu_0 \|v\|_{H^1}^2 = m_{\mathscr{E}}^2 \|v\|_{H^1}^2, \tag{3.2.2}$$

which establishes the energy semi-norm as a norm. Putting this together with (3.2.1) establishes the equivalence between the native and energy norms.

3.2.1 Linearized dual problems

Given a linear funcitonal $g \in L_2(\Omega)$, the objective in goal-oriented error estimation is to relate the residual to the error in the quantity of interest. This involves solving a dual problem whose solution z satisfies the relation $g(u - u_h) = \langle R(u_h), z \rangle$. In the linear case, the appropriate dual problem is the formal adjoint of the primal (cf. [31, 23]). For b nonlinear, the primal problem (3.1.2) does not have an exact formal adjoint. In this case we obtain the dual by linearization. Formally, given a numerical approximation u_j to the exact solution u, the residual is given by

$$R(u_j) := f - \mathscr{N}(u_j) = \mathscr{N}(u) - \mathscr{N}(u_j).$$

If $z^j \in H_0^1(\Omega)$ solves the following linearized dual problem

$$a(z^j, v) + \langle \mathscr{B}_j z^j, v\rangle = g(v), \quad \forall v \in H_0^1(\Omega), \tag{3.2.3}$$

where $g(v) := \int_\Omega gvdx$ and the operator $\mathscr{B}_j$ is given by

$$\mathscr{B}_j := \int_0^1 b'(\xi u + (1-\xi)u_j)\, d\xi = \int_0^1 b'(u_j + (u-u_j)\xi)\, d\xi, \tag{3.2.4}$$

then the goal-oriented error $g(e_j)$ can be represented exactly by the inner product of z^j and $R(u_j)$:

$$g(e_j) = \langle R(u_j), z^j\rangle.$$

In fact, by definition of the residual $R(u_j)$, we have

$$\langle z^j, R(u_j)\rangle = a(z^j, e_j) + \langle z^j, b(u) - b(u_j)\rangle = a(z^j, e_j) + \langle \mathscr{B}_j z^j, v\rangle = g(e_j).$$

Here we used the integral mean value identity:

$$b(u) - b(u_j) = \int_0^1 b'(u_j + (u-u_j)\xi)\, d\xi (u-u_j) = \mathscr{B}_j(u-u_j).$$

The derivation and numerical use of the linearized dual problem is further discussed in [15, 14, 21].

Unfortunately, the dual problem (3.2.3) is not practical because the operator $\mathscr{B}_j$ depends on the exact solution u. In order to define a computable dual operator, we introduce the approximate operator $b'(u_j)$, which lead to the following approximate dual problem: Find $\hat{z}^j \in H_0^1(\Omega)$ such that

$$a(\hat{z}^j, v) + \langle b'(u_j)\hat{z}^j, v\rangle = g(v), \quad \forall v \in H_0^1(\Omega). \tag{3.2.5}$$

The equation (3.2.5) is instrumental for defining a computable *a posteriori* error indicator for the dual problem.

A further difficulty arises in the analysis of the goal-oriented adaptive algorithm driven by the *a posteriori* error estimators for the approximate dual problem (3.2.5). Due to the dependence on u_j, (3.2.5) changes at each step of the adaptive algorithm. This is one of the essential differences of the nonlinear problem as compared to the linear cases in the previous literature (cf. [32, 23]). To handle this obstacle, we introduce the limiting dual problem: Find $\hat{z} \in H_0^1(\Omega)$ such that

$$a(\hat{z}, v) + \langle b'(u)\hat{z}, v\rangle = g(v), \quad \forall v \in H_0^1(\Omega). \tag{3.2.6}$$

While the operator $b'(u)$ is a function of the exact solution u and is not a computable quantity, it is the operator used in the limit of both the linearized dual (3.2.3) and approximate dual problems (3.2.5) as $u_j \to u$. Therefore, both the linearized and approximate sequences approach the same limiting problem (3.2.6). Our contraction result in Theorem 3.5.9 is written with respect to the limiting dual problem as defined by the operator $b'(u)$.

3.2.2 Finite Element Approximation

For a given conforming, shape-regular triangulation $\mathscr{T}$ of Ω consisting of close simplices $T \in \mathscr{T}$, we define the finite element space

$$\mathbb{V}_{\mathscr{T}} := H_0^1(\Omega) \cap \prod_{T \in \mathscr{T}} \mathbb{P}_n(T) \quad \text{and } \mathbb{V}_k := \mathbb{V}_{\mathscr{T}_k}, \tag{3.2.7}$$

where $\mathbb{P}_n(T)$ is the space of polynomials degree degree n over T. For any subset $\mathscr{S} \subseteq \mathscr{T}$,

$$\mathbb{V}_{\mathscr{T}}(\mathscr{S}) := H_0^1(\Omega) \cap \prod_{T \in \mathscr{S}} \mathbb{P}_n(T). \tag{3.2.8}$$

Given a triangulation $\mathscr{T}$, we denote $h_{\mathscr{T}} := \max_{T \in \mathscr{T}} h_T$ where $h_T := |T|^{1/d}$. In particular, we denote $h_0 := h_{\mathscr{T}_0}$ for an initial (conforming, shape-regular) triangulation $\mathscr{T}_0$ of Ω. Then the adaptive algorithm discussed below generates a nested sequence of conforming refinements $\{\mathscr{T}_k\}$, with $\mathscr{T}_k \geq \mathscr{T}_j$ for $k \geq j$ meaning that $\mathscr{T}_k$ is a conforming triangulation of Ω based on certain refinements of $\mathscr{T}_j$. With this notation, we also simply denote by $\mathbb{V}_k := \mathbb{V}_{\mathscr{T}_k}$ the finite element space defined on $\mathscr{T}_k$.

The finite element approximation of the primal problem (3.1.2) reads: Find $u_k \in$

$\mathbb{V}_k$ such that

$$a(u_k,v_k)+\langle b(u_k),v_k\rangle = f(v_k),\ v_k\in\mathbb{V}_k, \tag{3.2.9}$$

and the finite element approximation of (3.2.5) linearized about u_j is given by: Find $\hat{z}_k^j\in\mathbb{V}_k$ such that

$$a(\hat{z}_k^j,v_k)+\langle b'(u_j)\hat{z}_k^j,v_k\rangle = g(v_k) \quad \text{for all } v_k\in\mathbb{V}_k. \tag{3.2.10}$$

Finally, for the purpose of analysis, we require the discrete limiting dual problem (cf. (3.2.6)) given by: Find $\hat{z}_k\in\mathbb{V}_k$ such that

$$a(\hat{z}_k,v_k)+\langle b'(u)\hat{z}_k,v_k\rangle = g(v_k) \quad \text{for all } v_k\in\mathbb{V}_k. \tag{3.2.11}$$

Existence and uniqueness of solutions to the primal problems (3.1.2) and (3.2.9) follow from standard variational or fixed-point arguments as in [38] and [29]. For the dual problems (3.2.5)-(3.2.6) and (3.2.10)-(3.2.11) the existence and uniqueness of solutions follow from the standard Lax-Milgram Theorem as in [17], since we assumed that $b'(\xi)\geq 0$.

We make the following assumption on the a priori L_∞ bounds of the solutions to the primal problems (3.1.2) and (3.2.9):

Assumption 3.2.2 (*A priori bounds*)**.** *Let u and u_k be the solution to* (3.1.2) *and* (3.2.9)*, respectively. We assume that there are $u_-,u_+\in L_\infty$ which satisfy*

$$u_-(x)\leq u(x), u_k(x)\leq u_+(x) \textit{ for almost every } x\in\Omega. \tag{3.2.12}$$

Remark 3.2.3. *The L_∞ bound on u follows from the standard maximum principle, as discussed in [3, Theorem 2.4] and [24, Theorem 2.3]. There has been a lot of literature on the L_∞ bounds on the discrete solution, usually with additional angle condition of the triangulation (cf. [28, 26, 27, 24] and the references cited therein). On the other hand, if b satisfies the (sub)critical growth condition, as stated in [3, Assumption (A4)], then the L_∞ bounds on the discrete solution u_k are satisfied without angle conditions on the mesh, see [3] for more details.*

Assumption 3.2.1 together with Assumption 3.2.2 yield the following properties on the continuous and discrete solutions as summarized below.

Proposition 3.2.4. *Let the problem data satisfy Assumption 3.2.1 and Assumption 3.2.2. The following properties hold:*

1) *b is Lipschitz on $[u_-, u_+] \cap H_0^1(\Omega)$ for a.e. $x \in \Omega$ with constant B.*
2) *b' is Lipschitz on $[u_-, u_+] \cap H_0^1(\Omega)$ for a.e. $x \in \Omega$ with constant Θ.*
3) *Let $\hat{z}$ the solution to* (3.2.6)*, $\hat{z}_j^j$ the solution to* (3.2.10) *and let $\hat{z}_j$ the solution to* (3.2.11)*. Then there are $z_-, z_+ \in L_\infty$ which satisfy*

$$z_-(x) < \hat{z}(x), \hat{z}_j(x), \hat{z}_j^j(x) \leq z_+(x) \text{ for almost every } x \in \Omega,\ j \in \mathbb{N}. \tag{3.2.13}$$

3.3 Goal Oriented AFEM

In this section, we describe the goal oriented adaptive finite element method (GOAFEM), which is based on the standard AFEM algorithm:

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}\,.$$

Below, we explain each procedure.

Procedure SOLVE. The procedure SOLVE involves solving (3.2.9) for u_j, computing $b'(u_j)$ to form problem (3.2.10) and solving (3.2.10) for $\hat{z}_j^j$. In the analysis that follows, we assume for simplicity that the exact Galerkin solution is found on each mesh refinement. In practice the nonlinear problem (3.2.9) may be solved by a standard inexact Newton + multilevel algorithm as in [2]. The approximate dual problem (3.2.10) may be solved by any standard linear-time iterative method.

Procedure ESTIMATE. We use a fairly standard residual-based element-wise error estimator for both primal and approximate dual problems. Recall that the residual of the primal problem is given by $R(v) = f - \mathscr{N}(v)$ with $\mathscr{N}(v) = -\nabla \cdot (A\nabla v) + b(v)$. For the limiting and approximate dual problems, we define the local strong form by $\hat{\mathscr{L}}^*(v) := -\nabla \cdot (A\nabla v) + b'(u)(v)$, and $\hat{\mathscr{L}}_j^*(v) := -\nabla \cdot (A\nabla v) + b'(u_j)(v)$. The limiting and approximate dual residuals given respectively by

$$R^*(v) := g - \hat{\mathscr{L}}^*(v), \text{ and } \hat{R}^*_j(v) := g - \hat{\mathscr{L}}^*_j(v). \tag{3.3.1}$$

The *jump residual* for both the primal and linearized dual problems is:

$$J_T(v) := [\![[A\nabla v] \cdot n]\!]_{\partial T},$$

where $[\![\ \cdot \ ]\!]$ is given by $[\![\phi]\!]_{\partial T} := \lim_{t\to 0} \phi(x+tn) - \phi(x-tn)$and n is taken to be the appropriate outward normal defined on ∂T. On boundary edges σ_b we have $[\![[A\nabla v] \cdot n]\!]_{\sigma_b} \equiv 0$ so that $[\![[A\nabla v] \cdot n]\!]_{\partial T} = [\![[A\nabla v] \cdot n]\!]_{\partial T \cap \Omega}$. The error indicator for the primal problem (3.2.9) is given by

$$\eta^2_{\mathscr{T}}(v,T) := h_T^2 \|R(v)\|^2_{L_2(T)} + h_T \|J_T(v)\|^2_{L_2(\partial T)}, \quad v \in \mathbb{V}_{\mathscr{T}}. \tag{3.3.2}$$

Similarly, the dual error-indicator is given by the approximate residual

$$\zeta^2_{\mathscr{T},j}(w,T) := h_T^2 \|\hat{R}^*_j(w)\|^2_{L_2(T)} + h_T \|J_T(w)\|^2_{L_2(\partial T)}, \quad w \in \mathbb{V}_{\mathscr{T}}. \tag{3.3.3}$$

This dual indicator is defined in terms of the approximate dual operator $b'(u_j)$ as this is a computable quantity given an an approximation u_j. In addition, for purpose of analysis we define the limiting dual error-indicator by

$$\zeta^2_{\mathscr{T}}(w,T) := h_T^2 \|\hat{R}^*(w)\|^2_{L_2(T)} + h_T \|J_T(w)\|^2_{L_2(\partial T)}, \quad w \in \mathbb{V}_{\mathscr{T}}. \tag{3.3.4}$$

We remark that the limiting dual indicator as given by (3.3.4) is not computable. For any given subset $\mathscr{S} \subset \mathscr{T}$, the error estimators on $\mathscr{S}$ are given by the l_2 sum of error indicators over elements in the space.

$$\eta^2_{\mathscr{T}}(v,\mathscr{S}) := \sum_{T\in\mathscr{S}} \eta^2_{\mathscr{T}}(v,T), \quad v \in \mathbb{V}_{\mathscr{T}}.$$

The dual energy estimator is:

$$\zeta^2_{\mathscr{T},j}(w,\mathscr{S}) := \sum_{T\in\mathscr{S}} \zeta^2_{\mathscr{T},j}(w,T), \quad w \in \mathbb{V}_{\mathscr{T}},$$

and the limiting estimator

$$\zeta^2_{\mathscr{T}}(w,\mathscr{S}) := \sum_{T\in\mathscr{S}} \zeta^2_{\mathscr{T}}(w,T), \quad w \in \mathbb{V}_{\mathscr{T}}.$$

To simplify the notation, below we will omit "$\mathscr{S}$" in the above definitions if $\mathscr{S} = \mathscr{T}$ and we will use η_k to denote $\eta_{\mathscr{T}_k}$, and similarly use $\zeta_{k,\cdot}$ to denote $\zeta_{\mathscr{T}_k,\cdot}$.

As in [8] it is not difficult to verify that the indicators for the primal and approximate (respectively limiting) dual problems satisfy the monotonicity property for $v \in \mathbb{V}(\mathscr{T}_1)$ and $\mathscr{T}_2 \geq \mathscr{T}_1$

$$\eta_2(v,\mathscr{T}_2) \leq \eta_1(v,\mathscr{T}_1),\ \zeta_{2,j}(v,\mathscr{T}_2) \leq \zeta_{1,j}(v,\mathscr{T}_1) \ \text{ and } \ \zeta_2(v,\mathscr{T}_2) \leq \zeta_1(v,\mathscr{T}_1). \quad (3.3.5)$$

For an element $T \in \mathscr{T}_2 \cap \mathscr{T}_1$

$$\eta_2(v,T) = \eta_1(v,T),\ \zeta_{2,j}(v,T) = \zeta_{1,j}(v,T) \ \text{ and } \ \zeta_2(v,T) = \zeta_1(v,T). \quad (3.3.6)$$

Procedure MARK. The Dörfler marking strategy for the goal-oriented problem is based on the following steps as in [32]:

1) Given $\theta \in (0,1)$, mark sets for each of the primal and dual problems:

   - Mark a set $\mathscr{M}_p \subset \mathscr{T}_k$ such that

   $$\eta^2_k(u_k,\mathscr{M}_p) \geq \theta^2 \eta^2_k(u_k,\mathscr{T}_k). \quad (3.3.7)$$

   - Mark a set $\mathscr{M}_d \subset \mathscr{T}_k$ such that

   $$\zeta^2_{k,k}(\hat{z}^k_k,\mathscr{M}_d) \geq \theta^2 \zeta^2_{k,k}(\hat{z}^k_k,\mathscr{T}_k). \quad (3.3.8)$$

2) Let $\mathscr{M} = \mathscr{M}_p \cup \mathscr{M}_d$ the union of sets found for the primal and dual problems

respectively.

As in [23] the set $\mathscr{M}$ differs from that in [32], where the set of lesser cardinality between $\mathscr{M}_p$ and $\mathscr{M}_d$ is used. As seen in (3.3.8) the mesh is marked with respect to the dual indicators of the approximate-sequence solutions $\hat{z}_k^k$ as these are computable quantities. Sets $\mathscr{M}_p$ and $\mathscr{M}_d$ with optimal cardinality (up to a factor of 2) can be chosen in linear time by binning the elements rather than performing a full sort [32].

Procedure REFINE. The refinement (including the completion) is performed according to newest vertex bisection which was first proposed in [36]. It has been proved that the bisection procedure will preserve the shape-regularity of the initial triangulation $\mathscr{T}_0$. The complexity and other properties of this procedure are now well-understood (see for example [6] and the references cited therein), and will simply be exploited here.

3.4 Contraction for the primal problem

Here we discuss the contraction of the primal problem (3.1.2), recalling results from [25], [24] and [2]. The contraction argument relies on three main convergence results, namely quasi-orthogonality, error-estimator as upper bound on error and estimator reduction. We include the analogous results here for the limiting dual problem when they are identical or nearly identical.

3.4.1 Quasi-orthogonality

Orthogonality in the energy-norm $|||u-u_2|||^2 = |||u-u_1|||^2 - |||u_2-u_1|||^2$ does not generally hold in the semilinear problem. We rely on the weaker quasi-orthogonality result to establish contraction of AFEM (GOAFEM). The proof of the quasi-orthogonality relies on the following L_2-lifting property.

Lemma 3.4.1 (L_2-lifting)**.** *Let the problem data satisfy Assumption 3.2.1 and Assumption 3.2.2. Let u be the exact solution to* (3.1.2)*, and $u_1 \in \mathbb{V}_1$ the Galerkin solution to* (3.2.9)*. Let $w \in H^{1+s}(\Omega) \cap H_0^1(\Omega)$ for some $0 < s \leq 1$ be the solution to the dual* *problem: Find* $w \in H_0^1(\Omega)$ *such that*

$$a(w,v) + \langle \mathscr{B}_1 w, v\rangle = \langle u - u_1, v\rangle, \quad v \in H_0^1(\Omega), \tag{3.4.1}$$

where the operator $\mathscr{B}_1$ *is defined by* $\mathscr{B}_1 := \int_0^1 b'(\xi u + (1-\xi)u_1)\; d\xi$. *As in [9, 16, 1] we assume the regularity*

$$|w|_{H^{1+s}(\Omega)} \leq K_R \|u - u_1\|_{L_2(\Omega)} \tag{3.4.2}$$

based on the continuity of the diffusion coefficients a_{ij} *and* $\mathscr{B}_1 \in L_\infty(\Omega)$ *. Then*

$$\|u - u_1\|_{L_2} \leq C_* h_0^s |||u - u_1|||. \tag{3.4.3}$$

Proof. The proof follows the standard duality arguments in [1], [23] and [7], adapted for the semilinear problem. Let $\mathscr{I}^h : H_0^1(\Omega) \to \mathbb{V}_1$ be a quasi-interpolator, satisfying

$$\|w - \mathscr{I}^h w\|_{H^1} \leq C_{\mathscr{I}} h_{\mathscr{T}_1}^s |w|_{H^{1+s}} \tag{3.4.4}$$

$$\|w - \mathscr{I}^h w\|_{L_2} \leq \hat{C}_{\mathscr{I}} h_{\mathscr{T}_1}^{1+s} |w|_{H^{1+s}}. \tag{3.4.5}$$

as discussed in [1], [37] and [23].

Consider the linearized dual problem (3.4.1) with $v = u - u_1 \in H_0^1(\Omega)$ expressed in primal form

$$a(u - u_1, w) + \langle \mathscr{B}_1(u - u_1), w\rangle = \|u - u_1\|_{L_2}^2. \tag{3.4.6}$$

By Galerkin orthogonality, for $\mathscr{I}^h w \in \mathbb{V}_1$

$$a(u - u_1, \mathscr{I}^h w) + \langle \mathscr{B}_1(u - u_1), \mathscr{I}^h w\rangle = 0. \tag{3.4.7}$$

Subtracting (3.4.7) from (3.4.6)

$$a(u - u_1, w - \mathscr{I}^h w) + \langle b(u) - b(u_1), w - \mathscr{I}^h w\rangle = \|u - u_1\|_{L_2}^2. \tag{3.4.8}$$

Then by (3.2.1) continuity of $a(\cdot,\cdot)$, the Hölder inequality and Lipschitz continuity of

b (Proposition 3.2.4):

$$\|u-u_1\|_{L_2}^2 \leq M_{\mathscr{E}}\|u-u_1\|_{H^1}\|w-\mathscr{I}^h w\|_{H^1} + B\|u-u_1\|_{L_2}\|w-\mathscr{I}^h w\|_{L_2}. \tag{3.4.9}$$

By coercivity (3.2.2), interpolation estimate (3.4.4), and regularity (3.4.2) on the first term on the RHS of (3.4.9)

$$\begin{aligned} M_{\mathscr{E}}\|u-u_1\|_{H^1}\|w-\mathscr{I}^h w\|_{H^1} &\leq \frac{M_{\mathscr{E}}}{m_{\mathscr{E}}} C_{\mathscr{I}} h_0^s |||u-u_1||| \, |w|_{H^{1+s}} \\ &\leq \frac{M_{\mathscr{E}}}{m_{\mathscr{E}}} K_R C_{\mathscr{I}} h_0^s |||u-u_1||| \, \|u-u_1\|_{L_2}. \end{aligned} \tag{3.4.10}$$

For the second term of (3.4.9), apply (3.4.5) followed by (3.4.2) and coercivity to the interpolation error yielding

$$\begin{aligned} B\|u-u_1\|_{L_2}\|w-\mathscr{I}^h w\|_{L_2} &\leq B\hat{C}_{\mathscr{I}} h_0^{1+s}\|u-u_1\|_{L_2}|w|_{H^{1+s}} \\ &\leq K_R B\hat{C}_{\mathscr{I}} h_0^{1+s}\|u-u_1\|_{L_2}\|u-u_1\|_{L_2} \\ &\leq (m_{\mathscr{E}}^{-1} K_R B\hat{C}_{\mathscr{I}} h_0) h_0^s \|u-u_1\|_{L_2} |||u-u_1|||. \end{aligned} \tag{3.4.11}$$

Applying (3.4.10) and (3.4.11) to (3.4.9), we obtain

$$\|u-u_1\|_{L_2} \leq m_{\mathscr{E}}^{-1} K_R \left(M_{\mathscr{E}} C_{\mathscr{I}} + B\hat{C}_{\mathscr{I}} h_0\right) h_0^s |||u-u_1|||. \tag{3.4.12}$$

This completes the proof. □

Remark 3.4.2. *This proof yields the related result for $u_i \in \mathbb{V}_i$ the solution to* (3.2.9)*, $i = 1,2$*

$$\|u_2-u_1\|_{L_2} \leq C_* h_0^s |||u_2-u_1|||. \tag{3.4.13}$$

The proof of (3.4.13) *follows by replacing u by u_2 in 3.4.1. In particular, the dual problem* (3.4.1) *is replaced by: Find $w \in \mathbb{V}_2$ such that*

$$a(w,v) + \langle \mathscr{B}_{12} w, v\rangle = (u_2-u_1, v), \quad v \in \mathbb{V}_2, \tag{3.4.14}$$

where the operator $\mathscr{B}_{12} := \int_0^1 b'(\xi u_2 + (1-\xi)u_1)\, d\xi$.

Remark 3.4.3. *As the dual problem* (3.4.1) *changes at each iteration, so may the regularity constant as given by* (3.4.2) *as well as the interpolation constants as given by* (3.4.4) *and* (3.4.5)*. As such, the previous lemma shows a $C_{*,k}$ for $k = 1,2,\ldots$. As the algorithm is run finitely many times, we consolidate these $C_{*,k}$ into a single constant C_* for simplicity of presentation.*

Now we are in position to show the quasi-orthogonality.

Lemma 3.4.4 (Quasi-orthogonality)**.** *Let the problem data satisfy Assumption 3.2.1 and Assumption 3.2.2. Let $\mathscr{T}_1, \mathscr{T}_2$ be two conforming triangulation of Ω with $\mathscr{T}_2 \geq \mathscr{T}_1$. Let $u \in H_0^1(\Omega)$ the exact solution to* (3.1.2)*, $u_i \in \mathbb{V}_i$ the solution to* (3.2.9)*, $i = 1,2$. There exists a constant $C_* > 0$ depending on the problem data $\mathbf{D}$ and initial mesh $\mathscr{T}_0$, and a number $0 < s \leq 1$ related to the angles of $\partial\Omega$, such that if the meshsize h_0 of the initial mesh satisfies $\bar{\Lambda} := Bm_{\mathscr{E}}^{-1} C_* h_0^s < 1$, then*

$$|||u - u_2|||^2 \leq \Lambda |||u - \bar{v}|||^2 - |||u_2 - \bar{v}|||^2, \quad \forall \bar{v} \in \mathbb{V}_2, \tag{3.4.15}$$

and in particular for $\bar{v} = u_1 \in \mathbb{V}_1 \subset \mathbb{V}_2$

$$|||u - u_2|||^2 \leq \Lambda_G |||u - u_1|||^2 - |||u_2 - u_1|||^2, \tag{3.4.16}$$

where

$$\Lambda := (1 - Bm_{\mathscr{E}}^{-1} C_* h_0^s)^{-1} \ \text{ and } \ \Lambda_G := (1 - BC_*^2 h_0^{2s})^{-1}$$

and C_* is the constant from Lemma 3.4.1.

Proof. For any given $\bar{v} \in \mathbb{V}_2$, we have

$$|||u - u_2|||^2 = |||u - \bar{v}|||^2 - |||\bar{v} - u_2|||^2 + 2a(u - u_2, \bar{v} - u_2). \tag{3.4.17}$$

By Galerkin orthogonality

$$a(u - u_2, v) + \langle b(u) - b(u_2), v \rangle = 0 \text{ for all } v \in \mathbb{V}_2, \tag{3.4.18}$$

and taking $v = \bar{v} - u_2$ in (3.4.18), we have

$$\begin{aligned} 2a(u-u_2, \bar{v}-u_2) &\leq 2|\langle b(u) - b(u_2), \bar{v}-u_2\rangle| \\ &\leq 2B\|u-u_2\|_{L_2}\|\bar{v}-u_2\|_{L_2}. \end{aligned} \tag{3.4.19}$$

Here we used Hölder inequality and the Lipschitz property on b (cf. Proposition 3.2.4).

In the case of (3.4.15) applying L_2-lifting Lemma 3.4.1 to the first factor on the RHS and (3.2.2) coercivity to the second followed by Young's inequality, we obtain

$$\begin{aligned} 2B\|u-u_2\|_{L_2}\|\bar{v}-u_2\|_{L_2} &\leq 2Bm_{\mathscr{E}}^{-1}C_*h_0^s|||u-u_2|||\,|||\bar{v}-u_2||| \\ &\leq Bm_{\mathscr{E}}^{-1}C_*h_0^s|||u-u_2|||^2 + Bm_{\mathscr{E}}^{-1}C_*h_0^s|||\bar{v}-u_2|||^2. \end{aligned} \tag{3.4.20}$$

Applying (3.4.20) via (3.4.19) to (3.4.17)

$$(1-Bm_{\mathscr{E}}^{-1}C_*h_0^s)|||u-u_2|||^2 \leq |||u-\bar{v}|||^2 - (1-Bm_{\mathscr{E}}^{-1}C_*h_0^s)|||\bar{v}-u_2|||^2.$$

Assuming $\bar{\Lambda} := Bm_{\mathscr{E}}^{-1}C_*h_0^s < 1$, we have

$$|||u-u_2|||^2 \leq \Lambda|||u-\bar{v}|||^2 - |||\bar{v}-u_2|||^2 \tag{3.4.21}$$

with $\Lambda = (1-Bm_{\mathscr{E}}^{-1}C_*h_0^s)^{-1}$.

In the case of (3.4.16) applying L_2-lifting (Lemma 3.4.1) to each norm on the RHS of (3.4.19) by means of Remark 3.4.2 then applying Young's inequality

$$\begin{aligned} 2B\|u-u_2\|_{L_2}\|u_1-u_2\|_{L_2} &\leq 2Bh_0^{2s}C_*^2|||u-u_2|||\,|||u_1-u_2||| \\ &\leq Bh_0^{2s}C_*^2|||u-u_2|||^2 + BC_*^2h_0^{2s}|||u_1-u_2|||^2. \end{aligned} \tag{3.4.22}$$

Following the same procedure as above yields

$$|||u-u_2|||^2 \leq \Lambda_G|||u-u_1|||^2 - |||u_1-u_2|||^2 \tag{3.4.23}$$

with $\Lambda_G = (1-BC_*^2h_0^{2s})^{-1}$ with the weaker mesh assumption $\bar{\Lambda}_G := BC_*^2h_0^{2s} < 1$. □

We note that the second Galerkin orthogonality estimate (3.4.23) sharpens our results but is not essential to establishing them.

3.4.2 Error Estimator as Global Upper-bound

The second key result for the contraction of the primal problem is the error estimator as a global upper bound on the energy error, up to a global constant. The result for the semilinear problem is established in [25] and [2] with a clear generalization to the approximate dual sequence, also see [8] and [31] for the linear cases. The proof of this result follows from the general *a posteriori* error estimation framework developed in [39, 40].

Lemma 3.4.5 (Error estimator as global upper-bound)**.** *Let the problem data satisfy Assumption 3.2.1 and Assumption 3.2.2. Let $\mathscr{T}_k$ be a conforming refinement of $\mathscr{T}_0$. Let $u \in H_0^1(\Omega)$ and $u_k \in \mathbb{V}_k$ be the solutions to* (3.1.2) *and* (3.2.9)*, respectively. Similarly, let $\hat{z} \in H_0^1(\Omega)$ and $\hat{z}_k \in \mathbb{V}_k$ be the solutions to* (3.2.6) *and* (3.2.11)*, respectively. Then there is a global constant C_1 depending only on the problem data $\mathbf{D}$ and initial mesh $\mathscr{T}_0$ such that*

$$|||u - u_k||| \leq C_1 \eta_k(u_k) \tag{3.4.24}$$

and

$$|||\hat{z} - \hat{z}_k||| \leq C_1 \zeta_k(\hat{z}_k). \tag{3.4.25}$$

3.4.3 Estimator Reduction

The local Lipschitz property as in [25], analogous to the local perturbation property established in [8], is a key step in establishing estimator reduction leading to the contraction result. For any $T \in \mathscr{T}$, we denote

$$\omega_T := T \cup \{T' \in \mathscr{T} \mid T \cap T' \text{ is a true-hyperface of } T\}. \tag{3.4.26}$$

Here, for a d-simplex T, a true-hyperface is a $d-1$ sub-simplex of T, e.g., a face in 3D or an edge in 2D. We also define the data estimator on each element $T \in \mathscr{T}$ as

$$\eta_{\mathscr{T}}^2(\mathbf{D}, T) = h_T^2 \left(\|\mathrm{div} A\|_{L_\infty(T)}^2 + h_T^{-2} \|A\|_{L_\infty(\omega_T)}^2 + B^2 \right), \tag{3.4.27}$$

and denote $\eta_{\mathscr{T}}(\mathbf{D},\mathscr{S}) = \max_{T\in\mathscr{S}} \eta_{\mathscr{T}}(\mathbf{D},T)$ for any subset $\mathscr{S} \subseteq \mathscr{T}$. In particular, we denote by $\eta_0 := \eta_{\mathscr{T}_0}(\mathbf{D},\mathscr{T}_0)$ the data estimator on the initial mesh. As the grid is refined, the data estimator satisfies the monotonicity property for refinements $\mathscr{T}_2 \geq \mathscr{T}_1$ (cf. [8]):

$$\eta_2(\mathbf{D},\mathscr{T}_2) \leq \eta_1(\mathbf{D},\mathscr{T}_1). \tag{3.4.28}$$

Lemma 3.4.6 (Local Lipschitz Property)**.** *Let the problem data satisfy Assumption 3.2.1 and Assumption 3.2.2. Let $\mathscr{T}$ be a conforming refinement of $\mathscr{T}_0$. Then for all $T \in \mathscr{T}$ and for any $v,w \in \mathbb{V}_{\mathscr{T}}$*

$$|\eta_{\mathscr{T}}(v,T) - \eta_{\mathscr{T}}(w,T)| \leq \bar{\Lambda}_1 \eta_{\mathscr{T}}(\mathbf{D},T)\|v-w\|_{H^1(\omega_T)}. \tag{3.4.29}$$

The constant $\bar{\Lambda}_1 > 0$ depends on the dimension d and the initial mesh $\mathscr{T}_0$.

Proof. The proof follows those in [8] and [23]. We sketch the proof below. From (3.3.2)

$$\eta^2_{\mathscr{T}}(v,T) := h_T^2\|R(v)\|^2_{L_2(T)} + h_T\|J_T(v)\|^2_{L_2(\partial T)}, \quad v \in \mathbb{V}_{\mathscr{T}}. \tag{3.4.30}$$

Set $e = v - w$ and by definition of the residual, we get

$$\begin{aligned} R(v) &= f - \mathscr{N}(w+e) \\ &= f + \nabla\cdot(A\nabla w) - b(w) + \nabla\cdot(A\nabla e) - \int_0^1 b'(w+\xi e)\, d\xi e \\ &= R(w) + \mathscr{D}(e), \end{aligned}$$

where $\mathscr{D}(e) := \nabla\cdot(A\nabla e) - \int_0^1 b'(w+\xi e)\, d\xi e$. Using the generalized triangle-inequality

$$\sqrt{(a+b)^2 + (c+d)^2} \leq \sqrt{a^2+c^2} + b + d, \qquad \text{for } a,b,c,d > 0$$

and linearity of the jump residual we have

$$\begin{aligned} \eta_{\mathscr{T}}(v,T) &= \left(h_T^2\|R(w)+\mathscr{D}(e)\|^2_{L_2(T)} + h_T\|J(w)+J(e)\|^2_{L_2(\partial T)}\right)^{1/2} \\ &\leq \eta_{\mathscr{T}}(w,T) + h_T\|\mathscr{D}(e)\|_{L_2(T)} + h_T^{1/2}\|J(e)\|_{L_2(\partial T)}. \end{aligned} \tag{3.4.31}$$

For the second term of (3.4.31), by triangle inequality we obtain

$$\|\mathscr{D}(e)\|_{L_2(T)} \leq \|\nabla\cdot(A\nabla e)\|_{L_2(T)} + \left\| \int_0^1 b'(w+\xi e)\; d\xi e \right\|_{L_2(T)}. \tag{3.4.32}$$

By the inverse inequality, the diffusion term satisfies the bound

$$\begin{aligned}\|\nabla\cdot(A\nabla e)\|_{L_2(T)} &\leq \|\mathrm{div}A\cdot\nabla e\|_{L_2(T)} + \|A : D^2 e\|_{L_2(T)} \\ &\leq \left(\|\mathrm{div}A\|_{L_\infty(T)} + C_I h_T^{-1}\|A\|_{L_\infty(T)}\right)\|\nabla e\|_{L_2(T)},\end{aligned} \tag{3.4.33}$$

where $D^2 e$ is the Hessian of e. The second term in (3.4.32) is bounded by

$$\left\| \int_0^1 b'(w+\xi e)\; d\xi e \right\|_{L_2(T)} \leq B\|e\|_{L_2(T)}. \tag{3.4.34}$$

The jump term in (3.4.31) satisfies

$$\begin{aligned}\|J(e)\|_{L_2(\partial T)} &\leq 2(d+1)\; C_T h_T^{-1/2}\|A\|_{L_\infty(\omega_T)}\|\nabla e\|_{L_2(\omega_T)} \\ &= C_J h_T^{-1/2}\|A\|_{L_\infty(\omega_T)}\|\nabla e\|_{L_2(\omega_T)},\end{aligned} \tag{3.4.35}$$

where C_T depends on the shape-regularity of the triangulation. Putting together (3.4.31), (3.4.33), (3.4.34) and (3.4.35), we obtain

$$\begin{aligned}\eta_{\mathscr{T}}(v,T) &\leq \eta_{\mathscr{T}}(w,T) + h_T\left(\|\mathrm{div}A\|_{L_\infty(T)} + (C_I+C_J)h_T^{-1}\|A\|_{L_\infty(\omega_T)} + B\right)\|e\|_{H^1(\omega_T)} \\ &\leq \eta_{\mathscr{T}}(w,T) + C_{TOT}\,\eta_{\mathscr{T}}(\mathbf{D},T)\|v-w\|_{H^1(\omega_T)}.\end{aligned} \tag{3.4.36}$$

This completes the proof. □

The local perturbation property as demonstrated in Lemma 3.4.6 (respectively, Lemma 3.5.4 below) leads to estimator reduction, one of the three key ingredients for contraction of the both the primal and combined quasi-errors. This result holds for both the primal and limiting dual problems, whose proof can be found in [8, Corollary 2.4] or [23, Theorem 3.4].

Theorem 3.4.7 (Estimator reduction)**.** *Let the problem data satisfy Assumption 3.2.1 and Assumption 3.2.2. Let $\mathscr{T}_1$ be a conforming refinements of $\mathscr{T}_0$, $\mathscr{M} \subset \mathscr{T}_1$ be the* *marked set, and* $\mathscr{T}_2 = REFINE(\mathscr{T}_1, \mathscr{M})$. *Let*

$$\Lambda_1 := (d+2)\bar{\Lambda}_1^2 m_{\mathscr{E}}^{-2} \quad \text{and } \lambda := 1 - 2^{-1/d} > 0$$

with $\bar{\Lambda}_1$ *from Lemma 3.4.6 (local Lipschitz property). Then for any* $v_1 \in \mathbb{V}_1$ *and* $v_2 \in \mathbb{V}_2$ *and* $\delta > 0$

$$\eta_2^2(v_2, \mathscr{T}_2) \leq (1+\delta)\left\{\eta_1^2(v_1, \mathscr{T}_1) - \lambda\eta_1^2(v_1, \mathscr{M})\right\} + (1+\delta^{-1})\Lambda_1\eta_0^2 |||v_2 - v_1|||^2. \tag{3.4.37}$$

Analogously for the limiting dual problem

$$\zeta_2^2(v_2, \mathscr{T}_2) \leq (1+\delta)\left\{\zeta_1^2(v_1, \mathscr{T}_1) - \lambda\zeta_1^2(v_1, \mathscr{M})\right\} + (1+\delta^{-1})\Lambda_1\eta_0^2 |||v_2 - v_1|||^2. \tag{3.4.38}$$

The contraction of the primal (semilinear) problem is established in [25] and [22] based on Lemma 3.4.4, Lemma 3.4.5 and Theorem 3.4.7 as discussed above.

Theorem 3.4.8 (Contraction of the primal problem). *Let the problem data satisfy Assumption 3.2.1 and Assumption 3.2.2. Let u the solution to* (3.1.2). *Let* $\theta \in (0,1]$, *and let* $\{\mathscr{T}_j, \mathbb{V}_j, u_j\}_{j\geq 0}$ *be the sequence of meshes, finite element spaces and discrete solutions produced by GOAFEM. Then there exist constants* $\gamma_p > 0$ *and* $0 < \alpha < 1$, *depending on the initial mesh* $\mathscr{T}_0$ *and marking parameter* θ *such that*

$$|||u - u_{j+1}|||^2 + \gamma_p\eta_{j+1}^2 \leq \alpha^2\left(|||u - u_j|||^2 + \gamma_p\eta_j^2\right). \tag{3.4.39}$$

3.5 Contraction and Convergence of GOAFEM

In this section, we discuss the contraction and convergence of the GOAFEM described in §3.3. In particular, we show that the GOAFEM algorithm generates a sequence $\{\mathscr{T}_j, \mathbb{V}_j, u_j, \hat{z}_j\}_{j\geq 0}$ which contracts not only in the primal error as shown in §3.4, but also in a linear combination of the primal and limiting dual error. We emphasize that it would be difficult to derive convergence results in terms of problem (3.2.3) or (3.2.5), because at each refinement the problem is changing. So we show contraction in terms of the error in the limiting dual problem (3.2.6) as the target equation is fixed over the entire adaptive algorithm. Our approach to contraction in this section again relies on three main components: quasi-orthogonality, error-estimator as upper bound on error and estimator reduction. Here we discuss the relevant results for the limiting dual problem with an emphasis on those that differ significantly from the corresponding results for the primal problem. Note the limiting dual problem is not computable. We connect the error for the limiting dual problem to the computable quantities in the GOAFEM algorithm. For this purpose, we introduce Lemma 3.5.7, converting between limiting and approximate estimators in order to apply the Dörfler property to a computable quantity; and Lemma 3.5.8, bounding the discrete error between approximate and limiting dual solutions in terms of the primal error. We put these results together in Theorem 3.5.9 to establish the contraction of the combined quasi-error. Finally, the contraction of this form of the error is related to the error in the quantity of interest in Corollary 3.5.10.

3.5.1 Quasi-orthogonality for Limiting-dual Problem

Similar to the proof of the quasi-orthogonality for the primal problem, we make use of an L_2-lifting argument for the limiting-dual problem. Let $\hat{z} \in H_0^1(\Omega)$ and $\hat{z}_1 \in \mathbb{V}_1$ be the solutions to (3.2.6) and (3.2.11), respectively. We again use the duality argument, and introduce the problem: Find $y \in H_0^1(\Omega)$ such that

$$a(y,v) + \langle b'(u)y, v\rangle = (\hat{z} - \hat{z}_1, v) \text{ for all } v \in H_0^1(\Omega) \tag{3.5.1}$$

Then we have the following L_2-lifting result for the limiting-dual problem.

Lemma 3.5.1 (Limiting-dual L_2-lifting)**.** *Let the problem data satisfy Assumption 3.2.1 and Assumption 3.2.2. Let $\mathscr{T}_1$ be a conforming triangulation, and $\hat{z} \in H_0^1(\Omega)$ and $\hat{z}_1 \in \mathbb{V}_1$ be the solutions to* (3.2.6) *and* (3.2.11)*, respectively. Assume that the solution y to* (3.5.1) *belongs to $H^{1+s}(\Omega) \cap H_0^1(\Omega)$ for some $0 < s \leq 1$ such that*

$$|y|_{H^{1+s}(\Omega)} \leq \bar{K}_R \|\hat{z} - \hat{z}_1\|_{L_2(\Omega)}. \tag{3.5.2}$$

Then

$$\|\hat{z}-\hat{z}_1\| \leq \hat{C}_* h_0^s |||\hat{z}-\hat{z}_1|||. \tag{3.5.3}$$

Proof. The proof follows that of Lemma 3.4.1. As in (3.4.12), we obtain for the limiting dual estimate (3.5.3)

$$\|\hat{z}-\hat{z}_1\|_{L_2} \leq m_{\mathscr{E}}^{-1} \bar{K}_R \left(M_{\mathscr{E}} \bar{C}_{\mathscr{I}} + B\breve{C}_{\mathscr{I}} h_0 \right) h_0^s |||\hat{z}-\hat{z}_1|||. \tag{3.5.4}$$

□

Remark 3.5.2. *To apply L_2-lifting to the difference between two Galerkin solutions, $\hat{z}_2 \in \mathbb{V}_2$ and $\hat{z}_1 \in V_1$ where $\mathscr{T}_2 \geq \mathscr{T}_1$, we use the same proof with* (3.5.1) *replaced by the problem: Find $y \in \mathbb{V}_2$ such that*

$$a(y,v) + \langle b'(u)y, v\rangle = \langle \hat{z}_2 - \hat{z}_1, v\rangle \text{ for all } v \in \mathbb{V}_2. \tag{3.5.5}$$

By means of Lemma 3.5.1, we obtain the quasi-orthogonality for the limiting-dual problem.

Lemma 3.5.3 (Quasi-orthogonality for Limiting Dual Problem)**.** *Let the problem data satisfy Assumption 3.2.1, and $\mathscr{T}_1, \mathscr{T}_2$ be two conforming triangulations with $\mathscr{T}_2 \geq \mathscr{T}_1$. Let $\hat{z} \in H_0^1(\Omega)$ the solution to* (3.2.6) *and $\hat{z}_i \in \mathbb{V}_i$ the solution to* (3.2.11)*, $i = 1,2$. There exists a constant $\hat{C}_* > 0$ depending on the problem data $\mathbf{D}$ and initial mesh $\mathscr{T}_0$, and a number $0 < s \leq 1$ related to the regularity of* (3.5.1)*, such that for sufficiently small h_0 we have*

$$|||\hat{z}-\hat{z}_2|||^2 \leq \hat{\Lambda} |||\hat{z}-\bar{v}|||^2 - |||\hat{z}_2-\bar{v}|||^2, \quad \forall \bar{v} \in \mathbb{V}_2, \tag{3.5.6}$$

and in particular for $\bar{v} = \hat{z}_1$

$$|||\hat{z}-\hat{z}_2|||^2 \leq \hat{\Lambda}_G |||\hat{z}-\hat{z}_1|||^2 - |||\hat{z}_2-\hat{z}_1|||^2 \tag{3.5.7}$$

where

$$\hat{\Lambda} := (1 - Bm_{\mathscr{E}}^{-1} \hat{C}_* h_0^s)^{-1} \text{ and } \hat{\Lambda}_G := (1 - B\hat{C}_*^2 h_0^{2s})^{-1}$$

and $\hat{C}_*$ is the constant from Lemma 3.5.1.

Proof. The proof follows same arguments as in Lemma 3.4.4, except that in place of the inequality in (3.4.18) we have for the limiting dual problem

$$a(u-u_2,v)+\langle b'(u)(\hat{z}-\hat{z}_2),v\rangle = 0 \text{ for all } v\in\mathbb{V}_2, \tag{3.5.8}$$

yielding

$$2a(\hat{z}-\hat{z}_2,\bar{v}-\hat{z}_2)\le 2B\|\hat{z}-\hat{z}_2\|_{L_2}\|\bar{v}-\hat{z}_2\|_{L_2}, \tag{3.5.9}$$

as in (3.4.19). The rest of the proof is similar to Lemma 3.4.4, and we omit it here. □

3.5.2 Estimator Perturbations for Dual Sequence

As we have seen in Theorem 3.4.7, the local Lipschitz (local perturbation) property (cf. Lemma 3.4.6) plays a key role in deriving the estimator reduction property used to convert between estimators on different refinement levels in both the primal and limiting dual problems. The following lemma gives similar local Lipschitz properties for the approximate and limiting dual problems on a given refinement level.

Lemma 3.5.4 (Local Lipschitz Property for Dual Estimators)**.** *Let the problem data satisfy Assumption 3.2.1 and Assumption 3.2.2. Let $\mathscr{T}$ be a conforming refinement of $\mathscr{T}_0$. Then for all $T\in\mathscr{T}$ and for any $v,w\in\mathbb{V}_{\mathscr{T}}$, it holds that*

$$|\zeta_{\mathscr{T},j}(v,T)-\zeta_{\mathscr{T},j}(w,T)|\le\bar{\Lambda}_1\eta_{\mathscr{T}}(\mathbf{D},T)\|v-w\|_{H^1(\omega_T)}. \tag{3.5.10}$$

In particular, for the error indicator of the limiting dual problem we have

$$|\zeta_{\mathscr{T}}(v,T)-\zeta_{\mathscr{T}}(w,T)|\le\bar{\Lambda}_1\eta_{\mathscr{T}}(\mathbf{D},T)\|v-w\|_{H^1(\omega_T)}. \tag{3.5.11}$$

The constant $\bar{\Lambda}_1>0$ depends on the dimension d and the regularity of the initial mesh $\mathscr{T}_0$.

Proof. The proof follows those in [8], [23] and is nearly identical to Lemma 3.4.6. It is sketched here. To prove (3.5.10), by (3.3.3) we have

$$\zeta^2_{\mathscr{T},j}(v,T) := h_T^2 \|\hat{R}_j^*(v)\|^2_{L_2(T)} + h_T \|J_T(v)\|^2_{L_2(\partial T)}, \quad v \in \mathbb{V}_{\mathscr{T}}. \tag{3.5.12}$$

Setting $e = v - w$ and applying linearity to the definition of the dual residual as given by (3.3.1), we obtain

$$\hat{R}_j^*(v) = g + \hat{\mathscr{L}}_j^*(w+e) = \hat{R}_j^*(w) + \hat{\mathscr{L}}_j^*(e).$$

By the same reasoning as (3.4.31), we get

$$\zeta_{\mathscr{T},j}(v,T) \le \zeta_{\mathscr{T},j}(w,T) + h_T \|\hat{\mathscr{L}}_j^*(e)\|_{L_2(T)} + h_T^{1/2} \|J(e)\|_{L_2(\partial T)}. \tag{3.5.13}$$

The term $\hat{\mathscr{L}}_j^*$ (respectively $\hat{\mathscr{L}}^*$ for the limiting dual) in (3.5.13) satisfies the same bound as the analogous term $\mathscr{D}$ in (3.4.31) of Lemma 3.4.6. Hence the bounds (3.5.10) and (3.5.11) hold with the same constants as in (3.4.29). □

With the help of Lemma 3.5.4, we are able to derive the following corollary, which addresses the error induced by switching between error indicators corresponding to the approximate and limiting dual problems on a given element.

Corollary 3.5.5. *Let the problem data satisfy Assumption 3.2.1 and Assumption 3.2.2. Let $\mathscr{T}$ be a conforming refinement of $\mathscr{T}_0$, and u, u_j are the solutions to* (3.1.2) *and* (3.2.9) *problems, respectively. Let Θ and K_Z the constants given in Proposition 3.2.4. For all $T \in \mathscr{T}$ and for $v, w \in \mathbb{V}_{\mathscr{T}} \cap [z_-, z_+]$ the dual indicator on $\mathscr{T}$ satisfies*

$$|\zeta_{\mathscr{T},j}(v,T) - \zeta_{\mathscr{T},k}(w,T)| \le \bar{\Lambda}_1 \eta_{\mathscr{T}}(\mathbf{D},T) \|v - w\|_{H^1(\omega_T)} + \Theta K_Z h_T \|u_j - u_k\|_{L_2(T)}. \tag{3.5.14}$$

In particular, for $\mathscr{T} = \mathscr{T}_1$, we have for the limiting estimator

$$|\zeta_{1,1}(v,T) - \zeta_1(w,T)| \le \bar{\Lambda}_1 \eta_1(\mathbf{D},T) \|v - w\|_{H^1(\omega_T)} + \Theta K_Z h_T \|u - u_1\|_{L_2(T)}, \tag{3.5.15}$$

$$|\zeta_1(w,T) - \zeta_{1,1}(v,T)| \le \bar{\Lambda}_1 \eta_1(\mathbf{D},T) \|v - w\|_{H^1(\omega_T)} + \Theta K_Z h_T \|u - u_1\|_{L_2(T)}. \tag{3.5.16}$$

Proof. By the definition of the residuals for the approximate dual problems, for any

$w \in \mathbb{V}_{\mathscr{T}}$ we have

$$\begin{aligned}\hat{R}_j^*(w) &= g + \nabla \cdot (A\nabla w) + b'(u_k)w + \big(b'(u_j) - b'(u_k)\big)\, w \\ &= \hat{R}_k^*(w) + \big(b'(u_j) - b'(u_k)\big)\, w. \qquad (3.5.17)\end{aligned}$$

Using (3.5.17) in the definition of the dual indicator (3.3.3) and applying a generalized triangle inequality

$$\begin{aligned}\zeta_{\mathscr{T},j}(w,T) &= \Big(h_T^2 \|\hat{R}_k^*(w) + (b'(u_j) - b'(u_k))w\|_{L_2(T)}^2 + h_T \|J_T(w)\|_{L_2(\partial T)}^2\Big)^{1/2} \\ &\le \Big(h_T^2 \|\hat{R}_k^*(w)\|_{L_2(T)}^2 + h_T \|J_T(w)\|_{L_2(\partial T)}^2\Big)^{1/2} + h_T \|b'(u_j) - b'(u_k)w\|_{L_2(T)} \\ &\le \zeta_{\mathscr{T},k}(w,T) + \Theta K_Z h_T \|u_j - u_k\|_{L_2(T)}. \qquad (3.5.18)\end{aligned}$$

Applying (3.5.10) in Lemma 3.5.4 to the estimate (3.5.18), we obtain (3.5.14). □

As an immediate consequence of Corollary 3.5.5, we have the following results on the error induced by switching between dual estimators over a collection of elements on a given refinement level. This estimate plays a key role in the contraction argument below, as we apply it to switching between the estimator for the limiting dual and the computed error estimators for the approximate dual problems in the GOAFEM algorithm.

Corollary 3.5.6. *Let the hypotheses of Corollary 3.5.5 hold. Then for any subsets $\mathscr{M}_1, \mathscr{M}_2 \subseteq \mathscr{T}_1$ and arbitrary $\delta_1, \delta_2, \delta_A, \delta_B > 0$*

$$\begin{aligned}\zeta_1^2(v,\mathscr{M}_1) \ge{}& (1+\delta_1)^{-1}(1+\delta_A)^{-1}\zeta_{1,1}^2(w,\mathscr{M}_1) \\ &- (1+\delta_1)^{-1}\delta_A^{-1}\Theta^2 K_Z^2 h_0^2 \|u - u_1\|_{L_2}^2 - (d+2)\delta_1^{-1}\bar{\Lambda}_1^2 \eta_0^2 \|v - w\|_{H^1}^2 \qquad (3.5.19)\end{aligned}$$

$$\begin{aligned}\zeta_{1,1}^2(w,\mathscr{M}_2) \ge{}& (1+\delta_2)^{-1}(1+\delta_B)^{-1}\zeta_1^2(v,\mathscr{M}_2) \\ &- (1+\delta_2)^{-1}\delta_B^{-1}\Theta^2 K_Z^2 h_0^2 \|u - u_1\|_{L_2}^2 - (d+2)\delta_2^{-1}\bar{\Lambda}_1^2 \eta_0^2 \|v - w\|_{H^1}^2. \qquad (3.5.20)\end{aligned}$$

Proof. The conclusions follow by squaring inequality (3.5.15) (respectively (3.5.16)), applying Young's inequality twice, and then summing over element $T \in \mathscr{M}_1$ (respectively $T \in \mathscr{M}_2$). The H^1 norm is summed over all elements $T \in \mathscr{T}_1$ counting each element $d+2$ times, the maximum number of elements in each patch ω_T. □

3.5.3 Contraction of GOAFEM

The main contraction argument Theorem 3.5.9 follows after two more lemmas. The first combines a sequence of estimates to convert the non-computable limiting estimator for the dual problem to a computable quantity, apply the Dörfler property and then convert back. The second relates the difference between the Galerkin solutions of the limiting and approximate dual problems to the primal error. Motivated by estimator reduction for the limiting dual problem as in equation (3.4.38)

$$\zeta_2^2(\hat{z}_2, \mathscr{T}_2) \leq (1+\delta)\left\{\zeta_1^2(\hat{z}_1, \mathscr{T}_1) - \lambda \zeta_1^2(\hat{z}_1, \mathscr{M})\right\} + (1+\delta^{-1})\Lambda_1 \eta_0^2 |||\hat{z}_2 - \hat{z}_1|||^2 \tag{3.5.21}$$

the following lemma addresses the conversion between the limiting estimator $\zeta_1^2(\hat{z}_1, \mathscr{M})$ and and the computable estimator $\zeta_{1,1}^2(\hat{z}_1^1, \mathscr{M})$ necessary for marking the mesh for refinement.

Lemma 3.5.7. *Let the problem data satisfy Assumption 3.2.1 and Assumption 3.2.2. Let Θ and K_Z as given by Proposition 3.2.4, C_* as given by Lemma 3.4.1 and Λ_1 as given in Lemma 3.4.7. Let*

u the solution to (3.1.2), *u_1 the solution to* (3.2.9),
$\hat{z}$ the solution to (3.2.6), *$\hat{z}_1$ the solution to* (3.2.11) *$\hat{z}_1^1$ the solution to* (3.2.10).

Let $\zeta_{1,1}(\hat{z}_1^1, \mathscr{M})$ satisfy the Dörfler property for $\mathscr{M} \subset \mathscr{T}_1$: $\zeta_{1,1}^2(\hat{z}_1^1, \mathscr{M}) \geq \theta^2 \zeta_{1,1}^2(\hat{z}_1^1, \mathscr{T}_1)$.

Then for arbitrary $\delta_1, \delta_2, \delta_A, \delta_B > 0$ *there is a* $\delta_4 > 0$ *such that*

$$\begin{aligned}
-\zeta_1^2(\hat{z}_1, \mathscr{M}) \leq & -\frac{\beta\theta^2}{(1+\delta_4)}\zeta_1^2(\hat{z}_1, \mathscr{T}_1) - \frac{(1-\beta)\theta^2}{(1+\delta_4)C_1^2}|||\hat{z} - \hat{z}_1|||^2 \\
& + \left(\frac{\theta^2}{(1+\delta_A)(1+\delta_2)\delta_B} + \frac{1}{\delta_A}\right)\frac{\Theta^2 K_Z^2 C_*^2 h_0^{2(1+s)}}{(1+\delta_1)}|||u - u_1|||^2 \\
& + \left(\frac{\theta^2}{(1+\delta_1)(1+\delta_A)\delta_2} + \frac{1}{\delta_1}\right)\Lambda_1\eta_0^2(\mathbf{D}, \mathscr{T}_0)|||\hat{z}_1 - \hat{z}_1^1|||^2. \qquad (3.5.22)
\end{aligned}$$

Proof. From Corollary 3.5.6, L_2-lifting 3.4.1 and coercivity (3.2.2)

$$\begin{aligned}
-\zeta_1^2(\hat{z}_1, \mathscr{M}) \leq & -(1+\delta_1)^{-1}(1+\delta_A)^{-1}\zeta_{1,1}^2(\hat{z}_1^1, \mathscr{M}) \\
& + (1+\delta_1)^{-1}\delta_A^{-1}\Theta^2 K_Z^2 h_0^2\|u - u_1\|_{L_2}^2 + \delta_1^{-1}\bar{\Lambda}_1^2(d+2)\eta_0^2\|\hat{z}_1 - \hat{z}_1^1\|_{H^1}^2 \\
\leq & -(1+\delta_1)^{-1}(1+\delta_A)^{-1}\zeta_{1,1}^2(\hat{z}_1^1, \mathscr{M}) \\
& + (1+\delta_1)^{-1}\delta_A^{-1}\Theta^2 K_Z^2 C_*^2 h_0^{2(1+s)}|||u - u_1|||^2 + \delta_1^{-1}\Lambda_1\eta_0^2|||\hat{z}_1 - \hat{z}_1^1|||^2
\end{aligned} \qquad (3.5.23)$$

with $\Lambda_1 := \bar{\Lambda}_1^2(d+2)m_{\mathscr{E}}^{-2}$. The Dörfler property may be applied to the first term on the RHS of (3.5.23)

$$-\zeta_{1,1}^2(\hat{z}_1^1, \mathscr{M}) \leq -\theta^2\zeta_{1,1}^2(\hat{z}_1^1). \qquad (3.5.24)$$

Converting back to he limiting estimator by (3.5.20) in Corollary 3.5.6

$$\begin{aligned}
-\zeta_{1,1}^2(\hat{z}_1^1) \leq & -(1+\delta_2)^{-1}(1+\delta_B)^{-1}\zeta_1^2(\hat{z}_1, \mathscr{M}) \\
& + (1+\delta_2)^{-1}\delta_B^{-1}\Theta^2 K_Z^2 C_*^2 h_0^{2(1+s)}|||u - u_1|||^2 + \delta_2^{-1}\Lambda_1\eta_0^2|||\hat{z}_1 - \hat{z}_1^1|||^2.
\end{aligned} \qquad (3.5.25)$$

Define δ_4 by

$$(1+\delta_4) := (1+\delta_1)(1+\delta_2)(1+\delta_A)(1+\delta_B). \qquad (3.5.26)$$

Then by plugging (3.5.24) and (3.5.25) in the first term on the RHS of (3.5.23), we obtain

$$\begin{aligned} -\zeta_1^2(\hat{z}_1,\mathscr{M}) \le & -\theta^2(1+\delta_4)^{-1}\zeta_1^2(\hat{z}_1) + \big(\theta^2(1+\delta_A)^{-1}(1+\delta_2)^{-1}\delta_B^{-1}+\delta_A^{-1}\big) \\ & \times (1+\delta_1)^{-1}\Theta^2 K_Z^2 C_*^2 h_0^{2(1+s)} |||u-u_1|||^2 \\ & + \big(\theta^2(1+\delta_1)^{-1}(1+\delta_A)^{-1}\delta_2^{-1}+\delta_1^{-1}\big)\,\Lambda_1\eta_0^2 |||\hat{z}_1-\hat{z}_1^1|||^2. \end{aligned} \tag{3.5.27}$$

Finally, we split the first term on the RHS of (3.5.27) into two pieces for some $\beta \in (0,1)$, and apply the upper-bound estimate (3.4.25) in Lemma 3.4.5 to the second piece yielding

$$\begin{aligned} -\zeta_1^2(\hat{z}_1,\mathscr{M}) \le & -\beta\theta^2(1+\delta_4)^{-1}\zeta_1^2(\hat{z}_1) - (1-\beta)\theta^2(1+\delta_4)^{-1}C_1^{-2}|||\hat{z}-\hat{z}_1|||^2 \\ & + \big(\theta^2(1+\delta_A)^{-1}(1+\delta_2)^{-1}\delta_B^{-1}+\delta_A^{-1}\big)(1+\delta_1)^{-1}\Theta^2 K_Z^2 C_*^2 h_0^{2(1+s)}|||u-u_1|||^2 \\ & + \big(\theta^2(1+\delta_1)^{-1}(1+\delta_A)^{-1}\delta_2^{-1}+\delta_1^{-1}\big)\,\Lambda_1\eta_0^2|||\hat{z}_1-\hat{z}_1^1|||^2. \end{aligned}$$

This completes the proof. □

We may convert $|||\hat{z}_1-\hat{z}_1^1|||$ in the last term on the RHS of (3.5.22) to the error $|||u-u_1|||$ as stated in the following lemma.

Lemma 3.5.8. *Let the problem data satisfy Assumption 3.2.1 and Assumption 3.2.2. Let* Θ *and* K_Z *the constants given in Proposition 3.2.4 and* C_* *and* $\hat{C}_*$ *the constants given by Lemmas 3.4.1 and 3.5.1, respectively. Let*

u *the solution to* (3.1.2), $\qquad u_1$ *the solution to* (3.2.9),

$\hat{z}_1$ *the solution to* (3.2.11), $\qquad \hat{z}_1^1$ *the solution to* (3.2.10).

Then

$$|||\hat{z}_1-\hat{z}_1^1||| \le \Theta K_Z C_* \hat{C}_* h_0^{2s} |||u-u_1|||. \tag{3.5.28}$$

Proof. Recall that

$$\hat{z}_1 \text{ solves } a(\hat{z}_1,v)+\langle b'(u)\hat{z}_1,v\rangle = g(v), \text{ for all } v\in\mathbb{V}_1, \tag{3.5.29}$$

$$\hat{z}_1^1 \text{ solves } a(\hat{z}_1^1,v)+\langle b'(u_1)\hat{z}_1^1,v\rangle = g(v), \text{ for all } v\in\mathbb{V}_1. \tag{3.5.30}$$

Subtracting (3.5.30) from (3.5.29) and rearranging terms, we get

$$a(\hat{z}_1-\hat{z}_1^1,v)+\langle(b'(u)-b'(u_1))\hat{z}_1,v\rangle=\langle b'(u_1)(\hat{z}_1^1-\hat{z}_1),v\rangle,\ v\in\mathbb{V}_1. \tag{3.5.31}$$

In particular, for $v=\hat{z}_1-\hat{z}_1^1\in\mathbb{V}_1$ equation (3.5.31) yields

$$\begin{aligned}|||\hat{z}_1-\hat{z}_1^1|||^2&=-\langle(b'(u)-b'(u_1))\hat{z}_1,\hat{z}_1-\hat{z}_1^1\rangle-\langle b'(u_1)(\hat{z}_1-\hat{z}_1^1),\hat{z}_1-\hat{z}_1^1\rangle\\&\le-\langle(b'(u)-b'(u_1))\hat{z}_1,\hat{z}_1-\hat{z}_1^1\rangle\end{aligned} \tag{3.5.32}$$

where in the last inequality, we used the monotonicity assumption of b in Assumption (3.2.1). Now applying the Lipschitz property of b', the *a priori* L_∞ bounds on the dual solution $\hat{z}_1$ (cf. Proposition 3.2.4), and both primal and dual L_2 lifting in (3.5.32), we obtain

$$\begin{aligned}|||\hat{z}_1-\hat{z}_1^1|||^2&\le\Theta K_Z\|u-u_1\|_{L_2}\|\hat{z}_1-\hat{z}_1^1\|_{L_2}\\&\le\Theta K_Z C_*\hat{C}_*h_0^{2s}|||u-u_1|||\,|||\hat{z}_1-\hat{z}_1^1|||\end{aligned} \tag{3.5.33}$$

from which the result follows. □

Now we are in position to show the contraction of GOAFEM in terms of the combined quasi-error which is a linear combination of the energy errors and error estimators in primal and limiting dual problems.

Theorem 3.5.9 (Contraction of GOAFEM)**.** *Let the problem data satisfy Assumption 3.2.1 and Assumption 3.2.2. Let*

u *the solution to* (3.1.2),	u_j *the solution to* (3.2.9),
$\hat{z}$ *the solution to* (3.2.6),	$\hat{z}_j$ *the solution to* (3.2.11).

Let $\theta\in(0,1]$, *and let* $\{\mathscr{T}_j,\mathbb{V}_j\}_{j\ge0}$ *be the sequence of meshes and finite element spaces produced by GOAFEM. Let* $\gamma_p>0$ *as given by Theorem 3.4.8. Then for sufficient small* *mesh size h_0, there exist constants $\gamma > 0, \pi > 0$ and $\alpha_D \in (0,1)$ such that*

$$
\begin{aligned}
&|||\hat{z} - \hat{z}_2|||^2 + \gamma \zeta_2^2(\hat{z}_2) + \pi |||u - u_2|||^2 + \pi \gamma_p \eta_2^2(u_2) \\
&\leq \alpha_D^2 \left(|||\hat{z} - \hat{z}_1|||^2 + \gamma \zeta_1^2(\hat{z}_1) + \pi |||u - u_1|||^2 + \pi \gamma_p \eta_1^2(u_1) \right). \qquad (3.5.34)
\end{aligned}
$$

Proof. For simplicity, we denote $\eta_0 = \eta_0(\mathbf{D}, \mathscr{T}_0)$ and $\zeta_k(\hat{z}_k) = \zeta_k(\hat{z}_k, \mathscr{T}_k),\ k = 1,2$. By the estimator reduction for the limiting dual problem (3.4.38), for arbitrary $\delta > 0$ we have

$$
\zeta_2^2(\hat{z}_2) \leq (1+\delta) \left\{ \zeta_1^2(\hat{z}_1) - \lambda \zeta_1^2(\hat{z}_1, \mathscr{M}) \right\} + (1 + \delta^{-1}) \Lambda_1 \eta_0^2 |||\hat{z}_2 - \hat{z}_1|||^2, \qquad (3.5.35)
$$

where $\lambda = 1 - 2^{-1/d}$. Recall the quasi-orthogonality estimate in the limiting dual problem from Lemma 3.5.3

$$
|||\hat{z} - \hat{z}_2|||^2 \leq \hat{\Lambda}_G |||\hat{z} - \hat{z}_1|||^2 - |||\hat{z}_2 - \hat{z}_1|||^2. \qquad (3.5.36)
$$

Adding (3.5.36) to a positive multiple γ (to be determined) of (3.5.35) and applying the results of Lemmas 3.5.7 and 3.5.8 obtain

$$
\begin{aligned}
|||\hat{z} - \hat{z}_2|||^2 + \gamma \zeta_2^2(\hat{z}_2) \leq A |||\hat{z} - \hat{z}_1|||^2 + \gamma M \zeta_1^2(\hat{z}_1) + D |||u - u_1|||^2 \\
+ \left(\gamma (1 + \delta^{-1}) \Lambda_1 \eta_0^2 - 1 \right) |||\hat{z}_2 - \hat{z}_1|||^2. \qquad (3.5.37)
\end{aligned}
$$

We first set $\gamma := (1+\delta^{-1})^{-1} \Lambda_1^{-1} \eta_0^{-2}$ to eliminate the last term in (3.5.37). This yields

$$
|||\hat{z} - \hat{z}_2|||^2 + \gamma \zeta_2^2(\hat{z}_2) \leq A |||\hat{z} - \hat{z}_1|||^2 + \gamma M \zeta_1^2(\hat{z}_1) + D |||u - u_1|||^2, \qquad (3.5.38)
$$

where the coefficients A and M of (3.5.38) are given by

$$
\begin{aligned}
A &= \hat{\Lambda}_G - (1-\beta) \lambda \theta^2 \delta (1+\delta_4)^{-1} C_1^{-2} \Lambda_1^{-1} \eta_0^{-2} \qquad (3.5.39) \\
M &= (1+\delta)(1 - \beta \lambda \theta^2 (1+\delta_4)^{-1}) \qquad (3.5.40)
\end{aligned}
$$

where δ_4 satisfies $(1+\delta_4) := (1+\delta_1)(1+\delta_2)(1+\delta_A)(1+\delta_B)$ as was given in (3.5.26).

For contraction, we require $A < 1$ and $M < 1$ for the coefficients defined by

(3.5.39) and (3.5.40), that is, we need to choose a $\beta \in (0,1)$ such that

$$\frac{\delta}{1+\delta}\frac{1+\delta_4}{\lambda\theta^2} < \beta < 1 - \frac{(\hat{\Lambda}_G - 1)\Lambda_C}{\delta}\frac{1+\delta_4}{\lambda\theta^2}, \tag{3.5.41}$$

with $\Lambda_C := C_1^2\Lambda_1\eta_0^2$. To demonstrate the existence of such a β, set

$$\delta_4 = \delta = \frac{1}{2}\lambda\theta^2. \tag{3.5.42}$$

Then we require the mesh size h_0 sufficiently small, such that

$$\hat{\Lambda}_G < 1 + \frac{\lambda^2\theta^4}{2(2+\lambda\theta^2)\Lambda_C}, \tag{3.5.43}$$

for the given $\theta \in (0,1)$. Note the conditions (3.5.42) and (3.5.43) guarantee that the interval in (3.5.41) is nonempty, so there exists a β such that

$$\frac{1}{2} < \beta < 1 - \frac{(\hat{\Lambda}_G - 1)\Lambda_C}{\lambda\theta^2}\left(1 + \frac{2}{\lambda\theta^2}\right).$$

It remains to control the last term in (3.5.38). For simplicity, we assume $\delta_1 = \delta_2 = \delta_A = \delta_B =: \delta_C$. Then the coefficient D in (3.5.38) is given by

$$D = \delta\lambda\Theta^2 K_Z^2 C_*^2 h_0^{2s}\left(\frac{\theta^2 + (1+\delta_C)^2}{(1+\delta_C)^2\delta_C}\right)\left(\frac{h_0^2}{\Lambda_1\eta_0^2(1+\delta_C)} + \hat{C}_*^2 h_0^{2s}\right). \tag{3.5.44}$$

To control the primal error term with the coefficient D as given by (3.5.44), we add a positive multiple π (to be determined) of the primal contraction result (3.4.39) of Theorem 3.4.8 to (3.5.39) yieding

$$\begin{aligned} &|||\hat{z} - \hat{z}_2|||^2 + \gamma\zeta_2^2(\hat{z}_2) + \pi|||u - u_2|||^2 + \pi\gamma_p\eta_2^2(u_2) \\ &\le A|||\hat{z} - \hat{z}_1|||^2 + \gamma M\zeta_1^2(\hat{z}_1) + (D + \alpha^2\pi)|||u - u_1|||^2 + \alpha^2\pi\gamma_P\eta_1^2(u_1). \end{aligned} \tag{3.5.45}$$

We choose π to ensure $D + \alpha^2\pi < \pi$, namely,

$$\pi > \frac{D}{1-\alpha^2} \tag{3.5.46}$$

and set

$$\alpha_D^2 := \max\left\{A, M, \frac{D+\alpha^2\pi}{\pi}, \alpha^2\right\} < 1. \tag{3.5.47}$$

Then the combined quasi-error satisfies the contraction property (3.5.34). □

For simplicity, we denote by

$$\bar{Q}^2(u_j,\hat{z}_j) = |||\hat{z}-\hat{z}_j|||^2 + \gamma\zeta_j^2(\hat{z}_j) + \pi|||u-u_j|||^2 + \pi\gamma_p\eta_j^2(u_j)$$

the combined quasi-error in (3.5.34). The following corollary gives the contraction of the error in the goal function, which is determined by the contraction of the combined quasi-error.

Corollary 3.5.10. *Let the assumptions in Theorem 3.5.9 hold. Then the error in the goal function is controlled by a constant multiple of the square of the combined quasi-error, i.e.,*

$$|g(u)-g(u_j)| \le C\bar{Q}_j^2(u_j,\hat{z}_j) \le \alpha_D^{2j} C\bar{Q}_0^2(u_0,\hat{z}_0). \tag{3.5.48}$$

Proof. Choosing the test function $v = u - u_j$ in (3.2.6), and by linearity and Galerkin orthogonality for the primal problem, we obtain

$$\begin{aligned} g(u)-g(u_j) &= a(\hat{z},u) + \langle b'(u)\hat{z},u\rangle - a(\hat{z},u_j) - \langle b'(u)\hat{z},u_j\rangle \\ &= a(u-u_j,\hat{z}) + \langle b'(u)(u-u_j),\hat{z}\rangle \\ &= a(u-u_j,\hat{z}) + \langle \mathscr{B}_j(u-u_j),\hat{z}\rangle + \langle (b'(u)-\mathscr{B}_j)(u-u_j),\hat{z}\rangle \\ &= a(u-u_j,\hat{z}-\hat{z}_j) + \langle b(u)-b(u_j),\hat{z}-\hat{z}_j\rangle + \langle (b'(u)-\mathscr{B}_j)(u-u_j),\hat{z}\rangle. \end{aligned} \tag{3.5.49}$$

The third term in the last line of (3.5.49) represents the error induced by switching from (3.2.6) to (3.2.3). This term may be bounded in terms of the constants and L_∞ estimates in Proposition 3.2.4 and

$$\|b'(u)-\mathscr{B}_j\| = \left\| \int_0^1 b'(u) - b'\left(u_j + \xi(u-u_j)\right) d\xi \right\| \le \frac{\Theta}{2}\|u-u_j\|,$$

yielding

$$\begin{aligned}\langle (b'(u)-\mathscr{B}_j)(u-u_j),\hat{z}\rangle &\leq K_Z\|b'(u)-\mathscr{B}_j\|_{L_2}\|u-u_j\|_{L_2}\\ &\leq \frac{1}{2}\Theta K_Z\|u-u_j\|_{L_2}^2. \end{aligned}\tag{3.5.50}$$

Then by (3.5.49), (3.5.50), the Cauchy-Schwarz inequality and L_2-lifting as in Lemmas 3.4.1 and 3.5.1

$$\begin{aligned}|g(u)-g(u_j)| &\leq |||u-u_j||| \, |||\hat{z}-\hat{z}_j||| + B\|u-u_j\|_{L_2}\|\hat{z}-\hat{z}_j\|_{L_2} + \frac{1}{2}\Theta K_Z\|u-u_j\|_{L_2}^2\\ &\leq (1+BC_*\hat{C}_*h_0^{2s})|||u-u_j||| \, |||\hat{z}-\hat{z}_j||| + \frac{1}{2}\Theta K_Z C_*^2 h_0^{2s}|||u-u_j|||^2\\ &\leq \frac{1}{2}\left(1+(\Theta K_Z C_* + B\hat{C}_*)C_*h_0^{2s}\right)|||u-u_j|||^2 + \frac{1}{2}(1+BC_*\hat{C}_*h_0^{2s})|||\hat{z}-\hat{z}_j|||^2.\end{aligned}\tag{3.5.51}$$

Therefore the error in the goal function is bounded above by a constant multiple of the square of the combined quasi-error $\bar{Q}^2(u_j,\hat{z}_j)$. Thus (3.5.48) follows by the contraction result in Theorem 3.5.9. □

3.6 Numerical Experiments

Here we present some numerical experiments implemented using FETK [20], which is a fairly standard set of finite element modeling libraries for approximating the solutions to systems of nonlinear elliptic and parabolic equations. In these experiments, we consider $\Omega = [0,1]^2$ and try to solve the following model problem:

$$\mathscr{N}(u) := -\Delta u + 3u^3 = f, \tag{3.6.1}$$

with homogeneous Dirichlet boundary condition. Here the source function f is chosen such that the exact solution is given by

$$u(x,y) = \frac{\sin(\pi x)\sin(\pi y)}{2(x-0.5)^2+2(y-0.5)^2+10^{-3}}.$$

We solve the nonlinear problem using both the AFEM and the GOAFEM algorithms. Figure 3.1 shows a typical adaptive mesh as well as the solution of this semilinear equation.

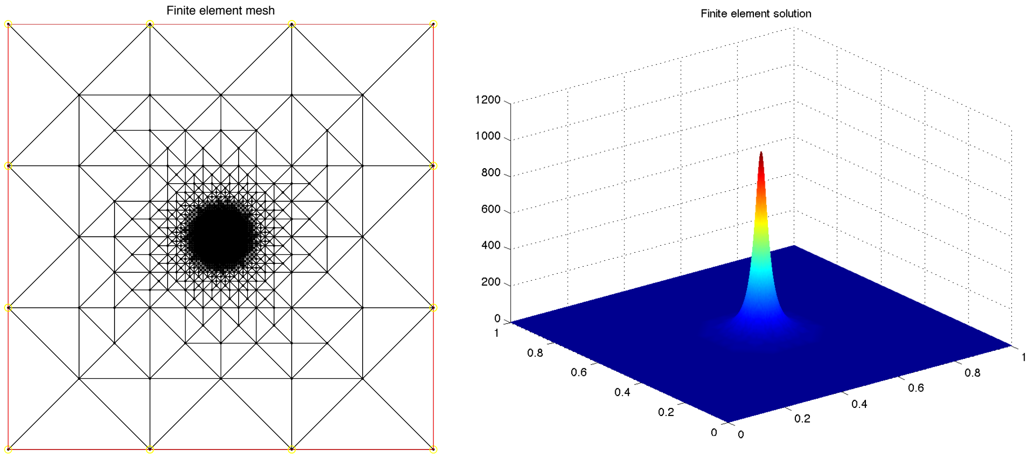

Figure 3.1: The mesh and finite element solution to the model problem after 12 GOAFEM iterations.

In the adaptive algorithms, we use the Dörfler marking strategies: (3.3.7) for the primal problem and (3.3.8) for the dual problem, with the same parameter $\theta = 0.4$. For the primal nonlinear problem (3.6.1), at each refinement we use a Newton-type iteration to solve the resulting nonlinear system of algebraic equations, which reduces the nonlinear residual to the tolerance $\|F(u)\|_{L_2} \leq 10^{-7}$. On the initial triangulation, we use a zero initial guess for the Newton iteration; then for each subsequent refinement, we interpolate the numerical solution from the previous step to the current triangulation and then use it as the initial guess for the Newton iteration. By doing this, we have a good initial guess for the Newton iteration so that one could expect a quadratic convergence rate of the nonlinear iterations. In fact, according to our numerical experiments, it usually takes 4 or 5 Newton iterations to reach the setting tolerance.

To test the performance of the GOAFEM algorithm, we take the goal function

$$g = 100e^{-100((x-0.5)^2+(y-0.5)^2)}, \tag{3.6.2}$$

so that it captures the singularity of the solution to (3.6.1). Thus the error in goal func-

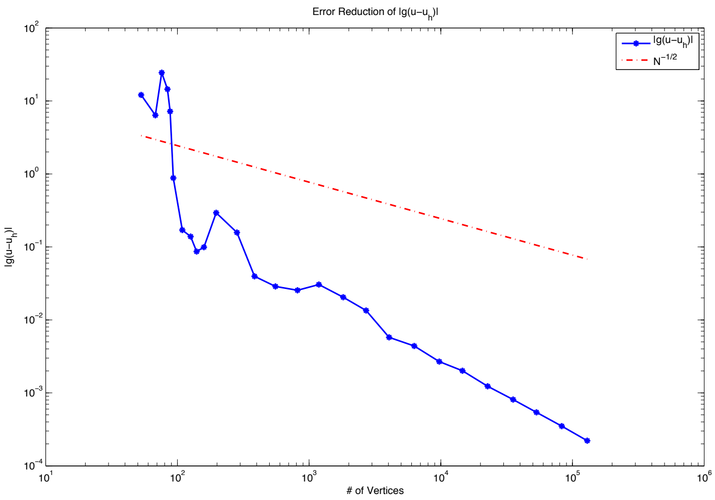

Figure 3.2: The reduction rate in the goal error $|g(u-u_h)|$.

tional $|g(u-u_h)|$ gives us an error by weighted average. Figure 3.2 shows the reduction rate of the goal error $|g(u-u_h)|$ for the GOAFEM algorithm. The oscillation at the first a few iterations in Figure 3.2 reflects the fact that the mesh size is not small enough, which is one of the requirements in our theory. But after a few iterations, the goal error reduces at certain rate, which was predicted by Theorem 3.5.48. This result confirms our theory.

For comparison with the standard AFEM algorithm, we also show in Figure 3.3 the error reduction in H^1 semi-norm $|u-u_h|_{H^1}$ for both AFEM and GOAFEM algorithms with the same goal function given in (3.6.2). As one can see from Figure 3.3, reducing H^1 error to the same magnitude with both algorithms, the GOAFEM algorithm takes about 5 fewer refinement steps than the standard AFEM algorithm.

3.7 Conclusion

In this article we developed convergence theory for a class of goal-oriented adaptive finite element algorithms for second order semilinear elliptic equations. We first introduced several approximate dual problems, and briefly discussed the target problem class. We then reviewed some standard facts concerning conforming finite element discretization and error-estimate-driven adaptive finite element methods (AFEM). We in-

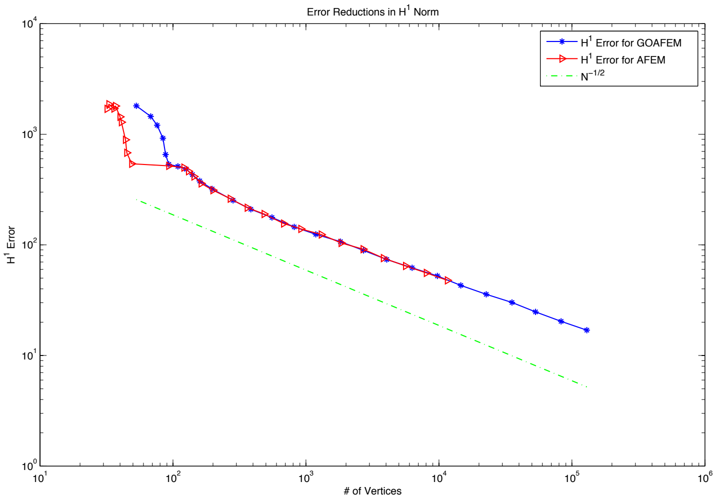

Figure 3.3: The reduction rate in the H^1 norm of the error $u - u_h$.

cluded a brief summary of *a priori* estimates for semilinear problems, and then described goal-oriented variations of the standard approach to AFEM (GOAFEM). Following the recent work of Mommer-Stevenson and Holst-Pollock for linear problems, we established contraction of GOAFEM for the primal problem. We also developed some additional estimates that make it possible to establish contraction of the combined quasi-error, and showed convergence in the sense of the quantity of interest. Some simple numerical experiments confirmed these theoretical predictions. Our analysis was based on the recent contraction frameworks for the semilinear problem developed by Holst, Tsogtgerel, and Zhu and Bank, Holst, Szypowski and Zhu and those for linear problems as in Cascon, Kreuzer, Nochetto and Siebert, and Nochetto, Siebert, and Veeser. In addressing the goal-oriented problem we based our approach on that of Mommer and Stevenson for symmetric linear problems and Holst and Pollock for nonsymmetric problems. However, unlike the linear case, we were faced with tracking linearized and approximate dual sequences in order to establish contraction with respect to the quantity of interest.

In the present paper we assume the primal and approximate dual solutions are solved on the same mesh at each iteration. The determination of strong convergence results for a method which solves the primal (nonlinear) problem on a coarse mesh and the dual on a fine mesh is the subject of future investigation.

Acknowledgments

MH was supported in part by NSF Awards 0715146 and 0915220, and by DOD/DTRA Award HDTRA-09-1-0036. SP was supported in part by NSF Award 0715146. YZ was supported in part by NSF Award 0715146 and by DOD/DTRA Award HDTRA-09-1-0036.

References

[1] O. Axelsson and V. A. Barker. *Finite element solution of boundary value problems: theory and computation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.

[2] R. Bank, M. Holst, R. Szypowski, and Y. Zhu. Convergence of AFEM for semilinear problems with inexact solvers, 2011.

[3] R. Bank, M. Holst, R. Szypowski, and Y. Zhu. Finite element error estimates for critical growth semilinear problems without angle conditions, 2011.

[4] R. Becker and R. Rannacher. A feed-back approach to error control in finite element methods: Basic analysis and examples. *East-West Journal of Numerical Mathematics*, 4:237–264, 1996.

[5] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numerica*, pages 1–102, 2001.

[6] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.

[7] S. Brenner and L. Scott. *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, third edition, 2008.

[8] J. M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.

[9] P. G. Ciarlet. *Finite Element Method for Elliptic Problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.

[10] W. Dahmen, A. Kunoth, and J. Vorloeper. *Convergence of Adaptive Wavelet Methods for Goal-oriented Error Estimation*. Sonderforschungsbereich 611, Singuläre Phänomene und Skalierung in Mathematischen Modellen. SFB 611, 2006.

[11] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM Journal on Numerical Analysis*, 33:1106–1124, 1996.

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

[13] D. Estep, M. Holst, and M. Larson. Generalized green's functions and the effective domain of influence. *SIAM J. Sci. Comput*, 26:1314–1339, 2002.

[14] D. Estep, M. Holst, and D. Mikulencak. Accounting for stability: A posteriori error estimates based on residuals and variational analysis. In *Communications in Numerical Methods in Engineering*, pages 200–2, 2001.

[15] D. Estep, M. G. Larson, and R. D. Williams. Estimating the error of numerical solutions of systems of reaction-diffusion equations. *Mem. Amer. Math. Soc.*, 146(696):101–109, 2000.

[16] L. C. Evans. *Partial Differential Equations (Graduate Studies in Mathematics, V. 19) GSM/19*. American Mathematical Society, 1998.

[17] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Springer-Verlag, 1977.

[18] M. Giles and E. Süli. Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality. *Acta Numerica*, 11:145–236, 2003.

[19] T. Grätsch and K.-J. Bathe. A posteriori error estimation techniques in practical finite element analysis. *Computers & Structures*, 83(4-5):235 – 265, 2005.

[20] M. Holst. Adaptive numerical treatment of elliptic systems on manifolds. 15(1–4):139–191, 2001. Available as arXiv:1001.1367 [math.NA].

[21] M. Holst. Applications of domain decomposition and partition of unity methods in physics and geometry, 2003.

[22] M. Holst, J. McCammon, Z. Yu, Y. Zhou, and Y. Zhu. Adaptive finite element modeling techniques for the Poisson-Boltzmann equation. *Communications in Computational Physics*, 11(1):179–214, 2012. Available as arXiv:1009.6034 [math.NA].

[23] M. Holst and S. Pollock. Convergence of goal oriented methods for nonsymmetric problems, 2011.

[24] M. Holst, R. Szypowski, and Y. Zhu. Two-grid methods for semilinear interface problems, 2012.

[25] M. Holst, G. Tsogtgerel, and Y. Zhu. Local and global convergence of adaptive methods for nonlinear partial differential equations, 2008.

[26] A. Jüngel and A. Unterreiter. Discrete minimum and maximum principles for finite element approximations of non-monotone elliptic equations. *Numer. Math.*, 99(3):485–508, 2005.

[27] J. Karatson and S. Korotov. Discrete maximum principles for finite element solutions of nonlinear elliptic problems with mixed boundary conditions. *Numerische Mathematik*, 99:669–698, 2005.

[28] T. Kerkhoven and J. W. Jerome. L_∞ stability of finite element approximations of elliptic gradient equations. *Numerische Mathematik*, 57:561–575, 1990.

[29] S. Kesavan. *Topics in Functional Analysis and Applications*. John Wiley and Sons, Inc., New York, NY, 1989.

[30] S. Korotov. A posteriori error estimation of goal-oriented quantities for elliptic type bvps. *Journal of Computational and Applied Mathematics*, 191(2):216 – 227, 2006.

[31] K. Mekchay and R. Nochetto. Convergence of adaptive finite element methods for general second order linear elliptic PDE. *SINUM*, 43(5):1803–1827, 2005.

[32] M. S. Mommer and R. Stevenson. A goal-oriented adaptive finite element method with convergence rates. *SIAM J. Numer. Anal.*, 47(2):861–886, 2009.

[33] K.-S. Moon, E. von Schwerin, A. Szepessy, and R. Tempone. Convergence rates for an adaptive dual weighted residual finite element algorithm. *BIT*, 46(2):367–407, 2006.

[34] R. H. Nochetto, K. G. Siebert, and A. Veeser. *Theory of adaptive finite element methods: an introduction*, pages 409 – 542. Springer, 2009.

[35] J. Oden and S. Prudhomme. Goal-oriented error estimation and adaptivity for the finite element method. *Computers and Mathematics with Applications*, 41:735–756, 2001.

[36] E. G. Sewell. Automatic generation of triangulations for piecewise polynomial approximation. In *Ph. D. dissertation*. Purdue Univ., West Lafayette, Ind., 1972.

[37] G. Strang and G. J. Fix. *An Analysis of the Finite Element Method*. Prentice-Hall (Series in Automatic Computation), Englewood Cliffs, N. J., 1973.

[38] M. Struwe. *Variational Methods*. Springer-Verlag, Berlin, Germany, 3 edition, 2000.

[39] R. Verfürth. A posteriori error estimates for nonlinear problems. finite element discretizations of elliptic equations. *Mathematics of Computation*, 62(206):445–475, Apr. 1994.

[40] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh refinement tecniques*. B. G. Teubner, 1996.

Chapter 3, in full, has been submitted for publication of the material as it may appear in SIAM Journal of Numerical Analysis, 2012, M. Holst, S. Pollock, Y. Zhu, SIAM, 2012. The dissertation author was the primary investigator and author of this paper.

Appendix A

Inverse Inequality

The inverse estimate shown here is a simplification of the one found in [2], Theorem 4.5.11. This estimate is used throughout many of the proofs in the second and third chapters, so we include this appendix to help the reader.

Let ω a bounded domain in $\mathbb{R}^d$ and $v \in \mathbb{V}(\omega)$, with $\dim(\mathbb{V}) < \infty$. For example, $\omega = T$ and $\mathbb{V} = \mathbb{V}_T$. Then there is a constant C, independent of $h_\omega = |\omega|^{1/d}$, with

$$|v|_{H^1(\omega)} \leq C h_\omega^{-1} \|v\|_{L_2(\omega)}.$$

Proof. Change variables to reference domain K with $\operatorname{diam}(K) = 1$. By the affine transformation $\xi = a + |\omega|^{-1} x$, define

$$K := \{a + |\omega|^{-1} x \big| \, x \in \omega\} = \{\xi(x) \big| \, x \in \omega\}. \tag{A.0.1}$$

Now define

$$\hat{v}(\xi) := v(x). \tag{A.0.2}$$

By the chain rule for each $j = 1, \dots, d$

$$\frac{\partial v}{\partial \xi_j} = \frac{\partial v}{x_j} \frac{\partial x_j}{\partial \xi_j} = |\omega|^{-1/d} \frac{\partial v}{x_j}, \text{ or more compactly } v_{,\xi_j} = |\omega|^{-1/d} v_{,x_j}. \tag{A.0.3}$$

Then $d\xi = d_{\xi_1} \cdots d_{\xi_d} = |\omega|^{-1} dx$, and applying (A.0.2) and (A.0.3) for each $j = 1, \ldots, d$

$$\int_{\xi \in K} |\hat{v}(\xi)_{,\xi_j}|^2 d\xi = \int_{x \in \omega} |\omega|^{2/d} |v(x)_{,x_j}|^2 |\omega|^{-1} dx = |\omega|^{(2-d)/d} \int_{x \in \omega} |v(x)_{,x_j}|^2 dx. \tag{A.0.4}$$

Summing (A.0.4) over squares of partial derivatives to obtain the gradient squared

$$\int_K |\nabla_\xi \hat{v}|^2 d\xi = |\omega|^{(2-d)/d} \int_\omega |\nabla_x v|^2 dx \iff |\hat{v}|^2_{H^1(K)} = |\omega|^{(2-d)/d} |v|^2_{H^1(\omega)}. \tag{A.0.5}$$

By the same change of variables,

$$\int_K |\hat{v}|^2 d\xi = |\omega|^{-1} \int_\omega |v|^2 dx \iff \|\hat{v}\|^2_{L_2(K)} = |\omega|^{-1} \|v\|^2_{L_2(\omega)}. \tag{A.0.6}$$

By equivalence of $H^1(K)$ and $L_2(K)$ norms over finite-dimensional spaces, there is a constant C with

$$|\hat{v}|^2_{H^1(K)} \leq \|\hat{v}\|^2_{H^1(K)} \leq C \|\hat{v}\|^2_{L_2(K)}. \tag{A.0.7}$$

By (A.0.5), (A.0.6) and (A.0.7), we have finally

$$|\omega|^{(2-d)/d} |v|^2_{H^1(\omega)} \leq C |\omega|^{-1} \|v\|^2_{L_2(\omega)} \iff |v|_{H_1(\omega)} \leq C |\omega|^{-1/d} \|v\|_{L_2(\omega)}. \tag{A.0.8}$$

□

A more general result may be obtained by applying technique this to derivatives of higher degree, as in [2].

Remark A.0.1. *This proof is an example of a common technique in finite element analysis, and in particular* a posteriori *error analysis, in which a change of variables is made to a reference domain where a property (here, equivalence of norms) is applied to bound a quantity modulo a constant C,* which may be a function of the domain. *Here we define our reference domain by an affine map from the element domain to satisfy* $|K| \equiv 1$. *By this method, we establish C as independent of the volume (or diameter) of T. The reference element $K = K(T)$ defined this way is not the same for all elements T, only its volume is the same. As such, the constant $C = C(K(T))$ is not necessarily the same for each element T, and it may contain factors relating to shape-regularity. These* *factors may be absorbed into a global constant C by merit of a shape regular mesh as produced by e.g., newest vertex bisection. Otherwise, the shape regularity factors could similarly be determined by a transformation from the barycentric coordinates of $K(T)$ (or T directly) to a* global *reference domain $\hat{K}$.*

Appendix B

Quasi-Interpolant Estimates

Here we establish the quasi-interpolant estimates given by (B.0.3) and (B.0.4). The proof follows the discussion of quasi-interpolants as given in [7]. These results are used in the proof of the estimator as a global upper bound on the energy error for the symmetric problem as in [8]. The discussion is included here as while the upper-bound estimate is a standard result, the details of this particular estimate are difficult to locate in the literature. Both the definition of the quasi-interpolant and the proof of the estimate are similar in form to the Scott-Zhang interpolant as given in [6] and [9]. The main difference between the two interpolants is the construction of the quasi-interpolant in terms of a dual basis over elements, as compared to a dual basis over element true-hyperfaces.

Let the mesh satisfy the following conditions:

1) The initial mesh $\mathscr{T}_0$ is conforming.

2) The mesh is refined by newest vertex bisection [1], [8] at each iteration.

Let the finite element spaces be given by

$$\mathbb{V}_{\mathscr{T}} := H_0^1(\Omega) \cap \prod_{T \in \mathscr{T}} \mathbb{P}_n(T) \quad \text{and } \mathbb{V}_k := \mathbb{V}_{\mathscr{T}_k}. \tag{B.0.1}$$

For subsets $\mathscr{S} \subseteq \mathscr{T}$, define

$$\mathbb{V}_{\mathscr{T}}(\mathscr{S}) := H_0^1(\Omega) \cap \prod_{T \in \mathscr{S}} \mathbb{P}_n(T), \tag{B.0.2}$$

where $\mathbb{P}_n(T)$ is the space of polynomials degree degree n over T. For a conforming mesh $\mathscr{T}_1$ with a conforming refinement $\mathscr{T}_2$, we say $\mathscr{T}_2 \geq \mathscr{T}_1$.

Lemma B.0.2. *Let $\mathscr{T}_2 \geq \mathscr{T}_1$ and $T \in \mathscr{T}_1$. For $v_q \in \mathbb{V}_1$ a quasi-interpolant of $v \in \mathbb{V}_2$, there is a global constant C such that*

$$\|v - v_q\|_{L_2(T)} \leq C h_T \|v\|_{H^1(\Omega_T)}, \tag{B.0.3}$$

$$\|v - v_q\|_{L_2(\partial T)} \leq C h_T^{1/2} \|v\|_{H^1(\Omega_T)}, \tag{B.0.4}$$

where

$$\Omega_T := \bigcup \{\tilde{T} \in \mathscr{T}_1 : \tilde{T} \cap T \neq \emptyset\}. \tag{B.0.5}$$

Proof of (B.0.3). First we will introduce the concept of a dual (to the Lagrange) basis, and establish some basic properties. We then use the dual basis to define the quasi-interpolant v_q and establish the estimate (B.0.3).

Let $N_T := \{x_i\}_{i=1}^M$ denote the set of nodes associated with element T with respect a basis of $\mathbb{P}_n(T)$. Denote the local nodal basis of $\mathbb{P}_n(T)$ by $\{\phi_{T,i} : x_i \in N_T\}$ where the basis functions satisfy the Lagrange property

$$\phi_{T,i}(y) = \delta_{x_i,y},\ y \in N_T. \tag{B.0.6}$$

Now consider a dual basis of $\mathbb{P}_n(T)$. Denote the basis functions by $\{\phi^*_{T,i} : x_i \in N_T\}$, where the $\phi^*_{T,i}$ are defined by the property

$$\langle \phi_{T,i}, \phi^*_{T,j} \rangle_{L_2(T)} = \delta_{ij}. \tag{B.0.7}$$

The set of nodes on element T is effectively defined on a reference element K with $|K| \equiv 1$, and mapped by affine transformation onto T as in (A.0.1). The reference nodal basis functions $\{\hat{\phi}_{K,\hat{v}}\}_{\hat{v} \in N(K)}$ are defined to satisfy the Lagrange property on the reference element, and mapped to the corresponding nodes on the element T. By the pointwise nature of the Lagrange property observe for $\hat{\phi}_{T,i}$ defined by $\hat{\phi}_{T,i}(\xi) := \phi_{T,i}(x)$

$$\hat{\phi}_{T,i} \equiv \phi_{K,i}. \tag{B.0.8}$$

In contrast, the dual basis on element T is scaled by $|T|$. In particular, by (B.0.7) and (B.0.8)

$$\delta_{ij} = \int_T \phi^*_{T,i}\phi_{T,j}\,dx = \int_K \hat{\phi}^*_{T_i}\hat{\phi}_{T,j}|T|\,d\xi = \int_K \left(|T|\hat{\phi}^*_{T,i}\right)\phi_{K,j}\,d\xi = \int_K \phi^*_{K,i}\phi_{K,j}\,d\xi, \tag{B.0.9}$$

from which we obtain the relation

$$\hat{\phi}^*_{T,i} = |T|^{-1}\phi^*_{K,i}. \tag{B.0.10}$$

From (B.0.10) obtain

$$\|\phi_{T,i}\|_{L_2(T)} = \left(|T|^2|T|^{-1}\int_K \phi_{K,i}\phi_{K,i}d\xi\right)^{1/2} = |T|^{-1/2}\|\phi_{K,i}\|_{L_2(K)}, \tag{B.0.11}$$

where $\|\phi_{K,i}\|_{L_2(K)}$ is independent of the measure of T. This establishes

$$\|\phi_{T,i}\|_{L_2(T)} \leq C_{K_1}|T|^{-1/2}, \tag{B.0.12}$$

where C_{K_1} is dependent on the regularity of the initial mesh (see Remark A.0.1), but independent of the meshsize. Before defining the quasi-interpolants, the following discussion establishes that a dual basis with the above properties is well-defined.

The dual basis is well-defined: The definition of the dual basis follows from Hilbert space properties of $\mathbb{P}_n(T)$ with L_2 inner-product $\langle\cdot,\cdot\rangle := \langle\cdot,\cdot\rangle_T$. Consider an arbitrary $\psi \in \mathbb{P}_n(T)$ and its expansion by basis functions

$$\psi = \sum_{x_j\in N_T} a_j\phi_{T,j}.$$

Then $f_i(\psi) = a_i$ is a bounded linear functional on $\mathbb{P}_n(T)$. By the Riesz-representation Theorem, there exists a unique element φ_i such that $\langle\varphi_i,\psi\rangle = f_i(\psi)$. By (B.0.7)

$$\langle\phi^*_{T,i},\psi\rangle = \int_T \sum_{x_j\in N_T}\phi^*_{T,i}a_j\phi_{T,j} = a_i,$$

we see $\phi^*_{T,i} \equiv \varphi_i$. This shows existence of the $M = \#N_T$ functions $\phi^*_{T,i}$, $i = 1\dots M$. To establish independence, suppose $\sum_{x_i \in N_T} \alpha_i \phi^*_{T,i} = 0$. Then given any $x_j \in N_T$,

$$0 = \int_T \sum_{x_i \in N_T} \alpha_i \phi^*_{T,i} \phi_{T,j} = \sum_{x_i \in N_T} \alpha_i \int_T \phi^*_{T,i} \phi_{T,j} = \alpha_j.$$

Equipped with a dual basis for each $T \in \mathscr{T}_1$, we now define the quasi-interpolant $v_q \in \mathbb{V}_1$. For any nodal point $x \in \left(\bigcup_{T \in \mathscr{T}_1} N_T\right) \cap \overset{\circ}{\Omega}$, we now select a $T_x \in \mathscr{T}_1$ with $x \in T_x$. For any nodal point x in the interior of some $T \in \mathscr{T}_1$ there is only one choice for T_x. However, if x lies on a true hyperface, there is a choice between two elements, and if x lies on a vertex of the triangulation, there is a choice between any element which contains x as a vertex. Now define v_q by its nodal values. For $x \in N_T$

$$\begin{aligned} v_q(x) &:= \int_{T_x} v \phi^*_{T_x,x}, && x \notin \partial\Omega \\ v_q(x) &:= 0, && x \in \partial\Omega. \end{aligned} \tag{B.0.13}$$

We now establish three key properties of quasi-interpolants.

Reproducing the value at nodes: For any $x_j \in \bigcup_{T \in \mathscr{T}_1} N_T$, $v_q(x_j) = v(x_j)$ when $T_x \in \mathscr{T}_2$.

Let $T_{x_j} \in \mathscr{T}_1 \cap \mathscr{T}_2$. Then $v\big|_{T_{x_j}} \in \mathbb{P}_m(T_{x_j})$, and may be written in terms of the basis $v\big|_{T_{x_j}} = \sum_{x_i \in T_{x_j}} \alpha_i \phi_{T_{x_j},i}\big|_{T_{x_j}}$. By the Lagrange property at nodes (B.0.6) $v(x_j) = \alpha_j$ and by (B.0.13), we have

$$v_q(x_j) = \int_{T_{x_j}} \left(\sum_{x_i \in T_{x_j}} \alpha_i \phi_{T_{x_j},i} \right) \phi^*_{T_{x_j},j} = \sum_{x_i \in T_{x_j}} \alpha_i \int_{T_{x_j}} \phi_{T_{x_j},i} \phi^*_{T_{x_j},j} = \alpha_j. \tag{B.0.14}$$

Define now the quasi-interpolant operator $Q_1 : \mathbb{V}_2 \to \mathbb{V}_1$ by

$$Q_1 v := v_q, \quad \text{for all } v \in \mathbb{V}_2. \tag{B.0.15}$$

Then Q_1 is a linear operator and in fact a projector. The property

$$Q_1 v_1 = v_1, \quad \text{for all } v_1 \in \mathbb{V}_1, \tag{B.0.16}$$

follows from the previous argument and the equivalence of polynomials which agree at nodal values. The linearity follows from the definition of nodal values (B.0.13), the linearity of the integral, and the quasi-interpolant as linear combination of basis functions.

L_2-norm bound: For any $T \in \mathscr{T}_1$ there is a constant C independent of the meshsize with

$$\|v_q\|_{L_2(T)} \leq C\|v\|_{L_2(\Omega_T)}. \tag{B.0.17}$$

If $\Omega_T \subseteq \mathscr{T}_1 \cap \mathscr{T}_2$, then by agreement on all the nodal values of T the polynomials must agree and

$$\|v_q\|_{L_2(T)} = \|v\|_{L_2(T)} \leq \|v\|_{L_2(\Omega_T)}. \tag{B.0.18}$$

Otherwise, write v_q as an expansion in nodal basis functions

$$v_q\big|_T = \sum_{x_i \in N_T} \phi_{T_{x_i},i}\big|_T v_q(x_i). \tag{B.0.19}$$

By property (B.0.12) and definition (B.0.13)

$$v_q(x_i) \leq C_{K_1}\|v\|_{L_2(\Omega_T)}|T_{x_i}|^{-1/2} \leq C_{K_1}\tilde{C}_\gamma\|v\|_{L_2(\Omega_T)}|T|^{-1/2}, \tag{B.0.20}$$

where $\tilde{C}_\gamma$ is a constant determined by the shape-regularity of the initial mesh and the property of newest vertex bisection refinement [1] that any two elements are separated by at most one generation. By (B.0.19) and (B.0.20)

$$\|v_q\|_{L_2(T)} \leq C_{K_1}\tilde{C}_\gamma\|v\|_{L_2(\Omega_T)}|T|^{-1/2} \sum_{x_i \in N_T} \|\phi_{T_{x_i},i}\|_{L_2(T)}. \tag{B.0.21}$$

By change of variables onto reference domain K,

$$\|\phi_{T_{x_i},i}\|_{L_2(T)} = |T|^{1/2}\|\hat{\phi}_{T_{x_i},i}\|_{L_2(K)}. \tag{B.0.22}$$

Combining (B.0.21) and (B.0.22) yields

$$\|v_q\|_{L_2(T)} \leq C_{K_1}C_\gamma C_\phi\|v\|_{L_2(\Omega_T)}, \tag{B.0.23}$$

where $C_\gamma = M\tilde{C}_\gamma$ and C_ϕ bounds the norm of basis functions on a global reference domain as in Remark A.0.1.

The interpolator reproduces any constant. If $v \in \mathbb{V}_2$ is constant then $v \in \mathbb{V}_1$ and by (B.0.16)

$$v_q = Q_1 v = v. \tag{B.0.24}$$

In particular, if $v\big|_T \in \mathbb{P}_0(T)$ then $v_q\big|_T = v\big|_T$.

With the above properties of the quasi-interpolant established, we now show

$$h_T^{-1}\|v - v_q\|_{L_2(T)} + |v - v_q|_{H^1(T)} \leq C_{K_5}|v|_{H^1(\Omega_T)}, \text{ for all } T \in \mathscr{T}_1. \tag{B.0.25}$$

First consider elements that do not touch the boundary of Ω. If $T \cap \partial\Omega = \emptyset$, then

$$h_T^{-1}\|v - v_q\|_{L_2(T)} + |v - v_q|_{H^1(T)} \leq C_{K_5}|v|_{H^1(\Omega_T)}. \tag{B.0.26}$$

Consider the operator $(I - Q_1)v = v - v_q$ for all $v \in \mathbb{V}_2$. By (B.0.24), $(I - Q_1)(v) = 0$ for all $v \in P_0(T)$. Applying the Bramble-Hilbert lemma [3] over reference domain K, by the change of variables as in (A.0.1)

$$|\hat{v} - \hat{v}_q| \leq C_{K_2}\|\widehat{I - Q_1}\|_{H^{-1}(K)}|\hat{v}|_{H^1(K)}. \tag{B.0.27}$$

Taking the $L_2(K)$ norm of both sides does not change the RHS as $|K| \equiv 1$

$$\|\hat{v} - \hat{v}_q\|_{L_2(K)} \leq C_{K_2}\|\widehat{I - Q_1}\|_{H^{-1}(K)}|\hat{v}|_{H^1(K)}. \tag{B.0.28}$$

Writing $C_{K_3} := \|\widehat{I - Q_1}\|_{H^{-1}(K)}$ where we may assume C_{K_3} a global constant by the same reasoning as Remark A.0.1

$$\|\hat{v} - \hat{v}_q\|_{L_2(K)} \leq C_{K_2}C_{K_3}|\hat{v}|_{H^1(K)}. \tag{B.0.29}$$

By equivalence of H^1- and L_2- norms on K, (B.0.29) yields

$$\|\hat{v} - \hat{v}_q\|_{L_2(K)} + |\hat{v} - \hat{v}_q|_{H_1(K)} \leq C_{K_2}C_{K_3}C_{K_4}|\hat{v}|_{H^1(K)}. \tag{B.0.30}$$

Changing variables back to T as in (A.0.6) and (A.0.7)

$$\|\cdot\|_{L_2(K)} = |T|^{-1/2}\|\cdot\|_{L_2(T)} \text{ and } |\cdot|_{H^1(K)} = h_T|T|^{-1/2}|\cdot|_{H^1(T)}. \tag{B.0.31}$$

Applying (B.0.31) to (B.0.30) yields the result (B.0.26).

In the second case $T \cap \partial\Omega \neq \emptyset$, so at least one of the $\tilde{T}$ that form Ω_T has a true hyperface on $\partial\Omega$. To handle this case, define another linear operator $\tilde{Q}_1 : \mathbb{V}_2 \to \mathbb{V}_1$, where for $x \in N_T$

$$\begin{cases} \tilde{Q}_1 v(x) := Q_1 v(x) & \text{if } x \notin \partial\Omega \\ \tilde{Q}_1 v(x) := v(x) = 0 & \text{if } x \in \partial\Omega. \end{cases} \tag{B.0.32}$$

As above, $\tilde{Q}_1$ is a linear operator which reproduces any constant. As any constant function in $\mathbb{V}_2$ must be zero, the second claim is trivial, and the first follows from the definition (B.0.32) and the linearity of Q_1. As such, $\tilde{Q}_1$ satisfies the hypotheses of the Bramble-Hilbert Lemma, and the second case follows as the first, establishing (B.0.25) for all $T \in \mathscr{T}_1$ from which the result (B.0.3) follows. □

Proof of (B.0.4). Applying the Trace theorem [4] on reference domain K

$$\|\hat{v} - \hat{v}_q\|_{L_2(\partial K)} \le C_{K_6}\|\hat{v} - \hat{v}_q\|_{H^1(K)}. \tag{B.0.33}$$

Converting (B.0.33) back to element T and multiplying through by a factor of h_T^{d-2}

$$\begin{aligned} h_T^{-1}\|v - v_q\|^2_{L_2(\partial T)} &= h_T^{d-2}\|\hat{v} - \hat{v}_q\|^2_{L_2(\partial K)} \\ &\le C^2_{K_6}\left(h_T^{-2}\|v - v_q\|^2_{L_2(T)} + |v - v_q|^2_{H^1(T)}\right). \end{aligned} \tag{B.0.34}$$

Combining (B.0.34) with (B.0.25)

$$h_T^{-1/2}\|v - v_q\|_{L_2(\partial T)} \le C_{K_5}C_{K_6}|v|_{H^1(\Omega_T)}, \tag{B.0.35}$$

from which the result (B.0.3) follows. □

Appendix C

Galerkin Method for Nonlinear Equations

Given a well-posed nonlinear problem we can't apply the Galerkin method as outlined in 1.1.4, as the first equality in (1.1.21) will not hold. Instead, we use a Newton iteration to a fixed tolerance as outlined below [5]. As an example, consider the problem discussed in Chapter 3, given by (1.1.5) with weak form (1.1.6). Let

$$F(u) := -\nabla \cdot A \nabla u + b(u) - f = 0, \tag{C.0.1}$$

yielding the weak-form equation

$$\langle F(u), v \rangle = a(u,v) + \langle b(u), v \rangle - f(v),\ \forall v \in H_0^1(\Omega). \tag{C.0.2}$$

We now have the problem equivalent to (1.1.6): Find $u \in H_0^1(\Omega)$ such that

$$\langle F(u), v \rangle = 0, \text{ for all } v \in H_0^1(\Omega). \tag{C.0.3}$$

The advantage of (C.0.3) is the problem is in suitable form to apply Newton's method. A basic Newton iteration has the form: Given $u^0 \in H_0^1(\Omega)$

$$\begin{aligned} \text{Solve for } h \in H_0^1(\Omega): \quad & \langle F'(u^k)h, v\rangle = -\langle F(u^k), v\rangle \ \text{ for all } v \in H_0^1(\Omega) && \text{(C.0.4)} \\ \text{Update:} \quad & u^{k+1} = u^k + h \\ \text{Stop if:} \quad & \|F(u)\| < tol. \end{aligned}$$

where tol is a predetermined tolerance. The term on the LHS of (C.0.4) is the Gâteaux derivative of F at u given by

$$\langle F'(u)w, v\rangle := \frac{d}{d\varepsilon}\langle F(u+\varepsilon w), v\rangle\big|_{\varepsilon=0}. \qquad \text{(C.0.5)}$$

For F as given by (C.0.2) expand the linear part by linearity and the nonlinear part by generalized Taylor expansion to obtain

$$\begin{aligned} \langle F'(u)w, v\rangle &= \frac{d}{d\varepsilon}\left[a(u+\varepsilon w, v) + \langle b(u+\varepsilon w), v\rangle - f(v)\right]_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon}\left[a(u,v) + \varepsilon a(w,v) + \langle b(u), v\rangle + \varepsilon\langle b'(u)w, v\rangle + \mathscr{O}(\varepsilon^2)\right]_{\varepsilon=0} \\ &= a(w,v) + \langle b'(u)w, v\rangle. && \text{(C.0.6)} \end{aligned}$$

Rewriting the iteration (C.0.4) in terms of (C.0.6): Given $u^0 \in H_0^1(\Omega)$

$$\begin{aligned} \text{Solve for } h \in H_0^1(\Omega): \quad & a(h,v) + \langle b'(u^k)h, v\rangle = -\langle F(u^k), v\rangle \ \text{ for all } v \in H_0^1(\Omega) \\ & && \text{(C.0.7)} \\ \text{Update:} \quad & u^{k+1} = u^k + h \\ \text{Stop if:} \quad & \|F(u)\| < tol. \end{aligned}$$

Finally, for the discrete problem as given by (1.1.22), given u_j^0 (for instance $u_j^0 := u_{j-1}$)

$$\begin{aligned} \text{Solve for } h \in \mathbb{V}_j(\Omega): \quad & a(h,v) + \langle b'(u_j^k)h, v\rangle = -\langle F(u_j^k), v\rangle \ \text{ for all } v \in \mathbb{V}_j && \text{(C.0.8)} \\ \text{Update:} \quad & u_j^{k+1} = u_j^k + h \\ \text{Stop if:} \quad & \|F(u_j)\| < tol. \end{aligned}$$

References

[1] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.

[2] S. Brenner and L. Scott. *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, third edition, 2008.

[3] P. G. Ciarlet. *Finite Element Method for Elliptic Problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.

[4] L. C. Evans. *Partial Differential Equations (Graduate Studies in Mathematics, V. 19) GSM/19*. American Mathematical Society, 1998.

[5] M. Holst. Course notes for math 273b, University of California San Diego, 2010.

[6] M. Holst and E. Titi. Determining projections and functionals for weak solutions of the Navier-Stokes equations. In *Recent Developments in Optimization Theory and Nonlinear Analysis, volume 204 of Contemporary Mathematics, Providence, Rhode Island*. American Mathematical Society, 1997.

[7] Y. Kondratyuk and R. Stevenson. An optimal adaptive finite element method for the Stokes problem. *SIAM J. Numer. Anal.*, 46(2):747–775, Feb. 2008.

[8] M. S. Mommer and R. Stevenson. A goal-oriented adaptive finite element method with convergence rates. *SIAM J. Numer. Anal.*, 47(2):861–886, 2009.

[9] R. H. Nochetto. Why adaptive finite element methods outperform classical ones. Hyderabad, India, 2010. Proceedings of the International Congress of Mathematicians.