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 then establish a contraction result for the primal problem. We then develop some additional estimates that make it possible to establish contraction of the combined quasi-error, and subsequently show convergence in the sense of the quantity of interest. Our analysis is 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 base our framework on that of Mommer and Stevenson for symmetric linear problems and Holst and Pollock for nonsymmetric problems. Unlike the linear case, one must track linearized and approximate dual sequences in order to establish contraction with respect to the quantity of interest.

CONTENTS

Date: March 8, 2012.

Key words and phrases. Adaptive methods, elliptic equations, semilinear problems, quasi-orthogonality, duality, approximation theory, residual-based error estimator, convergence, contraction, optimality, *a priori* estimates, *a posteriori* estimates, goal oriented.

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.

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

$$\begin{cases} -\nabla \cdot (A\nabla u) + b(u) = f, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega, \end{cases} \tag{1.1}$$

with $f \in L_2(\Omega)$ and $\Omega \subset \mathbb{R}^d$ a polyhedral domain, for $d = 2$ or 3. 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

$$a(u,v) + \langle b(u), v\rangle = f(v), \quad \forall v \in H_0^1(\Omega), \tag{1.2}$$

where

$$a(u,v) = \int_\Omega A\nabla u \cdot \nabla v \, dx. \tag{1.3}$$

In goal-oriented adaptive methods (cf. [10, 9] and the references therein for a detailed survey of these methods), one is are interested in a (usually linear) functional of the solution $g(u)$ rather than in the solution u itself. Our interest is in developing such an adaptive algorithm for semilinear problems along with a corresponding strong contraction result, following the recent approach in [19, 14] for linear problems. In particular, we develop a method for semilinear problems in which adaptive mesh refinement is driven both by residual-based approximations to the error in u, and in a sequence of approximate dual problems. While globally reducing the error in the primal problem necessarily yields a good approximation to the goal $g(u)$, 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 to $g(u)$ in fewer adaptive iterations, and hence fewer degrees of freedom.

Contraction for the semilinear problem is established in [16] and [2]. 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 problems by means of a computable sequence of approximate dual problems. We relate this result to a bound on the error in a goal-function, or quantity of interest. Following [16], 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
$$|\!|\!| u - u_2 |\!|\!|^2 \le \Lambda_G |\!|\!| u - u_1 |\!|\!|^2 - |\!|\!| u_2 - u_1 |\!|\!|^2.$$
2) Error estimator as upper bound on error: There exists $C_1 > 0$ such that
$$|\!|\!| u - u_k |\!|\!|^2 \le C_1 \eta_k^2(u_k, \mathcal{T}_k), \quad k = 1, 2.$$
3) Estimator reduction: For $\mathcal{M}$ the marked set that takes refinement $\mathcal{T}_1 \to \mathcal{T}_2$, for positive constants $\lambda < 1$ and Λ_1 and any $\delta > 0$
$$\eta_2^2(v_2, \mathcal{T}_2) \le (1+\delta)\{\eta_1^2(v_1, \mathcal{T}_1) - \lambda\eta_1^2(v_1, \mathcal{M})\} + (1+\delta^{-1})\Lambda_1\eta_0^2 |\!|\!| v_2 - v_1 |\!|\!|.$$

In the case of the primal problem, the mesh at each iteration may be marked for refinement with respect to the error indicators following the Dörfler marking strategy. In the case of the limiting dual problem, 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. 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*

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

which is the sum of the quasi-error as in [6] 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 5.11 establishes the contraction of the error in the goal function as shown in Corollary 5.12.

Our analysis is based on the recent contraction frameworks for the semilinear problem developed by Holst, Tsogtgerel, and Zhu [16] and Bank, Holst, Szypowski and Zhu [2] and those for linear problems as in Cascon, Kreuzer, Nochetto and Siebert [6], and Nochetto, Siebert, and Veeser [20]. In addressing the goal-oriented problem we base our framework on that of Mommer and Stevenson [19] for symmetric linear problems and Holst and Pollock [14] for nonsymmetric problems, and by combining these techniques we establish strong contraction of the method. The analysis of the goal-oriented method for nonlinear problems is signficantly more complex than the previous analysis for linear problems in [19, 14], where a much simpler analysis approach was possible. Here, we are faced with analyzing linearized and approximate dual sequences as well in order to establish contraction with respect to the quantity of interest. The linearized dual in the context of goal-oriented adaptive methods is described below, following e.g. Estep et. al in [10] and [8].

Outline of the paper. The remainder of the paper is structured as follows. In §2, we introduce the approximate, linear 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 §2.4 we include a brief summary of *a priori* estimates for the semilinear problem. In §3, we then describe a goal-oriented variation of the standard approach to AFEM (GOAFEM). In §4 we discuss contraction theorems for the primal problem. Lastly, in §5 we introduce additional estimates necessary for the contraction of the combined quasi-error and convergence in the sense of the quantity of interest.

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 several more practical and useful variants for computation and analysis.

2.1. **Linearized dual problem.**

Unlike the linear case as in [14] and [18], the primal problem does not have an exact formal adjoint. Instead, we consider the linearized dual problem as in [13], [10] and [8] associated to the average derivative of the nonlinear term.

The linearized dual operator $\mathcal{B}_j$ based on exact solution u and approximation u_j is given by

$$\mathcal{B}_j := \int_0^1 b'(\xi u + (1-\xi)u_j)\, d\xi = \int_0^1 b'(u_j + (u - u_j)\xi)\, d\xi. \tag{2.1}$$

By the integral mean value theorem [10] or equivalently a generalized Taylor expansion [13], the linearized dual operator satisfies the relation

$$\mathcal{B}_j(u - u_j) = b(u) - b(u_j). \tag{2.2}$$

In order to introduce a computable dual operator, one that is not a function of the exact solution u, we define the approximate dual operator $b'(u_j)$. This operator is instrumental for defining a computable *a posteriori* error indicator for the dual problem.

Our analysis also uses the limiting approximate dual operator $b'(u)$. While this operator 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 and approximate dual problems as $u_j \to u$. The contraction result Theorem 5.11 is written with respect to the limiting dual problem as defined by the operator $b'(u)$.

Consider the semilinear problem (1.2), where as in (1.3) we have

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

with $\langle\cdot,\cdot\rangle$ denoting the L_2 inner-product over $\Omega \subset \mathbb{R}^d$. The operators $\mathcal{B}_j,\ j = 1, 2, \ldots$ define a sequence of linearized dual problems: Find $z^j \in H_0^1(\Omega)$ such that

$$a(z^j, v) + \langle \mathcal{B}_j z^j, v\rangle = g(v), \quad \forall v \in H_0^1(\Omega). \tag{2.3}$$

Similarly, the operators $b'(u_j),\ j = 1, 2, \ldots$ define a sequence of approximate dual problems: Find $\hat{z}^j \in H_0^1(\Omega)$ such that

$$a(\hat{z}^j, v) + \langle b'(u_j)\hat{z}^j, v\rangle = g(v), \quad \forall v \in H_0^1(\Omega). \tag{2.4}$$

Both the linearized and approximate sequences approach the same limiting problem, find $\hat{z} \in H_0^1(\Omega)$ such that

$$a(\hat{z}, v) + \langle b'(u)\hat{z}, v\rangle = g(v), \quad \forall v \in H_0^1(\Omega). \tag{2.5}$$

Here, $a^*(\,\cdot\,,\,\cdot\,)$ the formal adjoint of $a(\,\cdot\,,\,\cdot\,)$, is equivalent to $a(\,\cdot\,,\,\cdot\,)$ for symmetric A. The goal functional is defined through

$$g(u) = \int_\Omega gu\, dx, \tag{2.6}$$

for given L_2 function $g\colon H_0^1(\Omega) \to \mathbb{R}$.

2.2. **Problem class, weak formulation, spaces and norms.** We will make the following assumptions on the data:

Assumption 2.1 (Problem data)**.** *The problem data $D = (A, b, f)$ and quantity of interest g satisfy*

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

$$\textit{ess inf}_{x\in\Omega}\lambda_{\textit{min}}(A(x)) = \mu_0 > 0, \tag{2.7}$$

$$\textit{ess sup}_{x\in\Omega}\lambda_{\textit{max}}(A(x)) = \mu_1 < \infty. \tag{2.8}$$

2) $b : \Omega \times \mathbb{R} \to \mathbb{R}$ *satisfies Assumption (A3) in* [2]*. For simplicity we write* $b(u)$ *instead of* $b(x, u)$*. Assume* b *monotone (increasing)*

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

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

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

$$\|v\|_{H^1}^2 = \langle \nabla v, \nabla v\rangle + \langle v, v\rangle. \tag{2.9}$$

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

$$a(u,v) \le \mu_1 \|u\|_{H^1}\|v\|_{H^1} = M_{\mathcal{E}}\|u\|_{H^1}\|v\|_{H^1}. \tag{2.10}$$

Define the energy semi-norm by the principal part of the differential operator

$$|\!|\!| v |\!|\!|^2 \coloneqq a(v,v). \tag{2.11}$$

Non-negativity follows from the Poincaré inequality with constant C_Ω

$$a(v,v) \ge \mu_0 |v|_{H^1}^2 \ge C_\Omega \mu_0 \|v\|_{H^1}^2 = m_{\mathcal{E}}^2 \|v\|_{H^1}^2, \tag{2.12}$$

which establishes the energy semi-norm as a norm. Putting this together with (2.10) establishes the equivalence between the native and energy norms.

2.3. **Finite Element Approximation.** We employ a standard conforming piecewise polynomial finite element approximation below.

Assumption 2.2 (Finite element mesh)**.** *We make the following assumptions on the underlying simplex mesh:*

1) *The initial mesh $\mathcal{T}_0$ is conforming.*
2) *The mesh is refined by newest vertex bisection* [4], [19] *at each iteration.*
3) *The initial mesh $\mathcal{T}_0$ is sufficiently fine. In particular, it satisfies (4.20).*

Based on assumptions 2.2 we have the following mesh constants.

1) Define
$$h_{\mathcal{T}} \coloneqq \max_{T\in\mathcal{T}} h_T, \qquad \text{where } h_T = |T|^{1/d}. \tag{2.13}$$
In particular, h_0 is the initial mesh diameter.
2) Define the mesh constant $\gamma_N = 2\gamma_r$ where
$$\gamma_r = \frac{h_0}{h_{min}} \quad \text{and} \quad h_{min} = \min_{T\in\mathcal{T}_0} h_T$$
then for any two elements $T, \tilde{T}$ in the same generation
$$h_T \le \gamma_r h_{\tilde{T}}$$
and as neighboring elements may differ by at most one generation for any two neighboring elements T and T'
$$h_T \le 2\gamma_r h_{T'} = \gamma_N h_{T'} \tag{2.14}$$
3) The minimal angle condition satisfied by newest vertex bisection implies the mesh-size h_T is comparable to h_σ, the size of any true-hyperface σ of T. In particular, there is a constant $\bar{\gamma}$
$$\frac{h_\sigma}{h_T} \le \bar{\gamma}^2 \text{ for all } T. \tag{2.15}$$

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

$$\mathbb{T}_N = \{\mathcal{T} \in \mathbb{T} \mid \#\mathcal{T} - \#\mathcal{T}_0 \le N\}.$$

For a conforming mesh $\mathcal{T}_1$ with a conforming refinement $\mathcal{T}_2$ we say $\mathcal{T}_2 \ge \mathcal{T}_1$.

Define the finite element space

$$\mathbb{V}_{\mathcal{T}} \coloneqq H_0^1(\Omega) \cap \prod_{T \in \mathcal{T}} \mathbb{P}_n(T) \quad \text{and } \mathbb{V}_k \coloneqq \mathbb{V}_{\mathcal{T}_k}. \tag{2.16}$$

For subsets $\mathcal{S} \subseteq \mathcal{T}$,

$$\mathbb{V}_{\mathcal{T}}(\mathcal{S}) \coloneqq H_0^1(\Omega) \cap \prod_{T \in \mathcal{S}} \mathbb{P}_n(T), \tag{2.17}$$

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

$$\omega_T \coloneqq T \cup \{T' \in \mathcal{T} \bigm| T \cap T' \text{ is a true-hyperface of } T\}. \tag{2.18}$$

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

$$a(u_k, v_k) + \langle b(u_k), v_k \rangle = f(v_k), \ v_k \in \mathbb{V}_k, \tag{2.19}$$

and the approximate dual problem linearized about u_j is given by: find $\hat{z}_k^j \in \mathbb{V}_k$ such that

$$a(\hat{z}_k^j, v_k) + \langle b'(u_j)\hat{z}_k^j, v_k \rangle = g(v_k) \quad \text{for all } v_k \in \mathbb{V}_k. \tag{2.20}$$

Finally, the discrete limiting dual problem is given by: find $\hat{z}_k \in \mathbb{V}_k$ such that

$$a(\hat{z}_k, v_k) + \langle b'(u)\hat{z}_k, v_k \rangle = g(v_k) \quad \text{for all } v_k \in \mathbb{V}_k. \tag{2.21}$$

Existence and uniqueness of solutions to the primal problems (1.2) and (2.19) follow from standard variational or fixed-point arguments as in [22] and [17]. For the dual problems (2.4) - (2.5) and (2.20) - (2.21) the result may be derived from the Lax-Milgram Theorem as in [12].

2.4. **A priori estimates.** We require the solutions to the primal and limiting and approximate dual problems satisfy L_∞ bounds. As discussed below, such bounds have been established assuming various additional conditions on either the nonlinearity b or on the angles of the mesh.

Assumption 2.3 (*A priori bounds*)**.** *Let u the solution to* (1.2)*, and u_j the solution to* (2.19)*. We assume the following L_∞ bounds on the primal and discrete primal solutions.*

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

$$u_-(x) < u(x), u_k(x) \le u_+(x) \text{ for almost every } x \in \Omega. \tag{2.22}$$

The L_∞ bound on u is discussed in [16] Lemma 7.9, [3] Theorem 2.4 and [15] Theorem 2.3 noting that Assumption 2.2 in [15] is a consequence of condition (2) of Assumption 2.1. The L_∞ bound on the discrete solution is demonstrated in [16] Lemma 7.9 and [15] Theorem 3.2 with the additional condition Assumption 3.1 of [15]. The L_∞ bound on the discrete solution u_k is also demonstrated without angle conditions on the mesh in [3] Corollary 4.4 . This case requires that the nonlinearity b satisfies the (sub)critical growth condition, as stated in [3] Assumption (A4).

Assumption 2.1 together with Assumption 2.3 yield the following properties as summarized below.

Proposition 2.4. *Let the problem data satisfy Assumption 2.1 and Assumption 2.3. 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 the mesh satisfy conditions (1) and (2) of Assumption 2.2. Let $\hat{z}$ the solution to (2.5), $\hat{z}_j^j$ *the solution to* (2.20) *and* $\hat{z}_j$ *the solution to* (2.21)*. Then there are* $z_-, z_+ \in L_\infty$ *which satisfy*

$$z_-(x) < \hat{z}(x), \hat{z}_j(x), \hat{z}_j^j(x) \leq z_+(x) \text{ for almost every } x \in \Omega,\ j \in \mathbb{N} \tag{2.23}$$

and there is a constant $K_Z := \max\{\|z_-\|_{L_\infty}, \|z_+\|_{L_\infty}\}$.

Remark 2.5. *The L_∞ bounds on the dual solutions as in (1) of Proposition* (2.4) *follow from the maximum principle as in* [12] *and L_∞ error estimates as in* [3].

3. Goal Oriented AFEM

The goal oriented adaptive finite element method (GOAFEM) is based on the standard AFEM algorithm:

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

Procedure SOLVE. The procedure SOLVE involves solving (2.19) for u_j, computing $b'(u_j)$ to form problem (2.20) and solving (2.20) for $\hat{z}_j^j$. In the analysis that follows, we suppose for simplicity the exact Galerkin solution is found on each mesh refinement.

In practice the nonlinear problem (2.19) may be solved by a standard inexact Newton + multilevel algorithm as in [2]. The approximate dual problem (2.20) may be solved by any standard linear-time iterative method so that the Galerkin solution to each problem is found up to a given tolerance. Convergence of the goal-oriented method assuming an inexact solution to the primal problem is currently under investigation by the present authors.

Procedure ESTIMATE. The estimation of the error on each element is determined by a fairly standard residual-based estimator, which we will now define. The *local strong form* of the nonlinear operator is

$$\mathcal{N}(v) := \nabla \cdot (A \nabla v) - b(v); \tag{3.1}$$

The *residual* for the primal problem, following the sign convention in [6]:

$$R(v) := f + \mathcal{N}(v). \tag{3.2}$$

For the limiting and approximate dual problems, we define the local strong form by

$$\hat{\mathcal{L}}^*(v) := \nabla \cdot (A \nabla v) - b'(u)(v), \text{ and } \hat{\mathcal{L}}_j^*(v) := \nabla \cdot (A \nabla v) - b'(u_j)(v). \tag{3.3}$$

The limiting and approximate dual residuals given respectively by

$$R^*(v) := g + \hat{\mathcal{L}}^*(v), \text{ and } \hat{R}_j^*(v) := g + \hat{\mathcal{L}}_j^*(v). \tag{3.4}$$

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

$$J_T(v) := [\![[A \nabla v] \cdot n]\!]_{\partial T} \tag{3.5}$$

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

$$[\![\phi]\!]_{\partial T} := \lim_{t \to 0} \phi(x + tn) - \phi(x - tn) \tag{3.6}$$

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

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

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

$$\eta_{\mathcal{T}}^2(v,T) := h_T^2\|R(v)\|_{L_2(T)}^2 + h_T\|J_T(v)\|_{L_2(\partial T)}^2, \quad v \in \mathbb{V}_{\mathcal{T}}. \tag{3.7}$$

The dual error-indicator is then given by the approximate residual

$$\zeta_{\mathcal{T},j}^2(w,T) := h_T^2\|\hat{R}_j^*(w)\|_{L_2(T)}^2 + h_T\|J_T(w)\|_{L_2(\partial T)}^2, \quad w \in \mathbb{V}_{\mathcal{T}}, \tag{3.8}$$

and the limiting dual error-indicator by

$$\zeta_{\mathcal{T}}^2(w,T) := h_T^2\|\hat{R}^*(w)\|_{L_2(T)}^2 + h_T\|J_T(w)\|_{L_2(\partial T)}^2, \quad w \in \mathbb{V}_{\mathcal{T}}. \tag{3.9}$$

The 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. The limiting dual indicator as given by (3.9) is not computable, but remains useful in the analysis. The error estimators are given by the l_2 sum of error indicators over elements in the space.

$$\eta_{\mathcal{T}}^2(v) := \sum_{T\in\mathcal{T}} \eta_{\mathcal{T}}^2(v,T), \quad v \in \mathbb{V}_{\mathcal{T}}. \tag{3.10}$$

The dual energy estimator is:

$$\zeta_{\mathcal{T},j}^2(w) := \sum_{T\in\mathcal{T}} \zeta_{\mathcal{T},j}^2(w,T), \quad w \in \mathbb{V}_{\mathcal{T}}, \tag{3.11}$$

and the limiting estimator

$$\zeta_{\mathcal{T}}^2(w) := \sum_{T\in\mathcal{T}} \zeta_{\mathcal{T}}^2(w,T), \quad w \in \mathbb{V}_{\mathcal{T}}. \tag{3.12}$$

To simplify the notation, welow we will use η_k to denote $\eta_{\mathcal{T}_k}$, and similarly use $\zeta_{k,\cdot}$ to denote $\zeta_{\mathcal{T}_k,\cdot}$.

As in [6] the indicators for the primal and approximate (respectively limiting) dual problems satisfy the monotonicity property for $v \in \mathcal{T}_1$ and $\mathcal{T}_2 \geq \mathcal{T}_1$

$$\eta_2(v,\mathcal{T}_2) \leq \eta_1(v,\mathcal{T}_1),\ \zeta_{2,j}(v,\mathcal{T}_2) \leq \zeta_{1,j}(v,\mathcal{T}_1) \ \text{ and } \ \zeta_2(v,\mathcal{T}_2) \leq \zeta_1(v,\mathcal{T}_1). \tag{3.13}$$

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

$$\eta_2(v,T) = \eta_1(v,T),\ \zeta_{2,j}(v,T) = \zeta_{1,j}(v,T) \ \text{ and } \ \zeta_2(v,T) = \zeta_1(v,T). \tag{3.14}$$

The data estimator over the mesh $\mathcal{T}$ or a subset $\mathcal{T}' \subset \mathcal{T}$ is given by the maximum data estimator over elements in the mesh or subset: For $\mathcal{T}' \subseteq \mathcal{T}$

$$\eta_{\mathcal{T}}(D,\mathcal{T}') = \max_{T\in\mathcal{T}'} \eta_{\mathcal{T}}(D,T).$$

The data estimator on the initial mesh

$$\eta_0 := \eta_{\mathcal{T}_0}(D,\mathcal{T}_0).$$

As the grid is refined, the data estimator satisfies the monotonicity property [6] for refinements $\mathcal{T}_2 \geq \mathcal{T}_1$

$$\eta_2(D,\mathcal{T}_2) \leq \eta_1(D,\mathcal{T}_1). \tag{3.15}$$

Procedure MARK. The Dörfler marking strategy for the goal-oriented problem is based on the following steps as in [19]:

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

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

$$\sum_{T\in\mathcal{M}_p} \eta_k^2(u_k,T) \geq \theta^2\eta_k^2(u_k,\mathcal{T}_k) \tag{3.16}$$

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

$$\sum_{T\in\mathcal{M}_d} \zeta_{k,k}^2(\hat{z}_k^k, T) \geq \theta^2 \zeta_{k,k}^2(\hat{z}_k^k, \mathcal{T}_k) \tag{3.17}$$

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

As in [14] the set $\mathcal{M}$ differs from that in [19], where the set of lesser cardinality between $\mathcal{M}_p$ and $\mathcal{M}_d$ is used. As seen in (3.17) the mesh is marked with respect to the dual indicators of the approximate-sequence solutions $\hat{z}_k^k$ as these are computable quantities. In the case of the semilinear problem the error reduced at each iteration is the combined quasi-error, a linear combination of energy error and estimators of the primal and limiting dual problems. This combined error is seen to contract based on the refinement satisfying the Dörfler property in terms of the primal and corresponding approximate dual problems. As such, the mesh is refined to satisfy the Dörfler property in each. Sets $\mathcal{M}_p$ and $\mathcal{M}_d$ 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 [19].

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.

Procedure REFINE. The refinement (including the completion) is performed according to newest vertex bisection [4]. The complexity and other properties of this procedure are now well-understood, and will simply be exploited here.

4. CONTRACTION FOR THE PRIMAL PROBLEM

Here we discuss the contraction of the primal problem (1.2), recalling results from [16], [15] 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 .

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 following is a variation on the quasi-orthogonality discussion in [16] and is related to the version for nonsymmetric linear problems as in [18] and [14]. The quasi-orthogonality proof relies on L_2-lifting, a fairly standard result included here for completeness. Here we show for $\bar{v} \in \mathbb{V}_2 \geq \mathbb{V}_1$

$$|||u-u_2|||^2 \leq \Lambda |||u-\bar{v}|||^2 - |||u_2-\bar{v}|||^2,$$

and in particular for $u_1 \in \mathbb{V}_1 \subset \mathbb{V}_2$

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

Lemma 4.1 (L_2-lifting)**.** *Let the problem data satisfy Assumption 2.1 and Assumption 2.3 and the mesh satisfy Assumption 2.2. Let u the variational solution to* (1.2)*, and $u_1 \in \mathbb{V}_1$ the Galerkin solution to* (2.19)*. Let B the constant given in Proposition 2.4. Assume for any $g \in L_2(\Omega)$ the solution w to the linearized dual problem* (2.3) *with $\mathcal{B}_1$ as given by* (4.3) *belongs to $H^{1+s}(\Omega) \cap H_0^1(\Omega)$ for some $0 < s < 1$ and*

$$|w|_{H^{1+s}(\Omega)} \leq K_R \|g\|_{L_2(\Omega)} \tag{4.1}$$

then

$$\|u - u_1\|_{L_2} \le C_* h_0^s \|\!|u - u_1|\!\|. \tag{4.2}$$

As discussed in [7] and [11] and [1] the regularity assumptions are reasonable based on the continuity of the diffusion coefficients a_{ij} and $\mathcal{B}_1 \in L_\infty(\Omega)$ where as in (2.1) the linearized dual operator with respect to u_1

$$\mathcal{B}_1 = \int_0^1 b'(\xi u + (1-\xi)u_1)\, d\xi = \int_0^1 b'(u_1 + (u - u_1)\xi)\, d\xi. \tag{4.3}$$

Proof. The proof follows the duality arguments in [1], [14] and [5], adapted for the semi-linear problem.

Let $w \in H_0^1(\Omega)$ the solution to the dual problem

$$a(w, v) + \langle \mathcal{B}_1 w, v\rangle = \langle u - u_1, v\rangle, \quad v \in H_0^1(\Omega). \tag{4.4}$$

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

$$\|w - \mathcal{I}^h w\|_{H^1} \le C_{\mathcal{I}} h_{\mathcal{T}_1}^s |w|_{H^{1+s}} \tag{4.5}$$

$$\|w - \mathcal{I}^h w\|_{L_2} \le \hat{C}_{\mathcal{I}} h_{\mathcal{T}_1}^{1+s} |w|_{H^{1+s}}. \tag{4.6}$$

as discussed in [1], [21] and [14].

Consider the linearized dual problem (4.4) with $v = u - u_1 \in H_0^1(\Omega)$ expressed in primal form

$$a(u - u_1, w) + \langle \mathcal{B}_1(u - u_1), w\rangle = \|u - u_1\|_{L_2}^2. \tag{4.7}$$

By Galerkin orthogonality, for $\mathcal{I}^h w \in \mathbb{V}_1$

$$a(u - u_1, \mathcal{I}^h w) + \langle \mathcal{B}_1(u - u_1), \mathcal{I}^h w\rangle = 0. \tag{4.8}$$

Subtracting (4.8) from (4.7)

$$a(u - u_1, w - \mathcal{I}^h w) + \langle \mathcal{B}_1(u - u_1), w - \mathcal{I}^h w\rangle = \|u - u_1\|_{L_2}^2. \tag{4.9}$$

Then by (2.10) continuity of $a(\,\cdot\,,\,\cdot\,)$, the relation (2.2), the Hölder inequality and Lipschitz continuity of b

$$\|u - u_1\|_{L_2}^2 \le M_{\mathcal{E}} \|u - u_1\|_{H^1} \|w - \mathcal{I}^h w\|_{H^1} + B\|u - u_1\|_{L_2} \|w - \mathcal{I}^h w\|_{L_2}. \tag{4.10}$$

By coercivity (2.12), interpolation estimate (4.5), and regularity (4.1) on the first term on the RHS of (4.10)

$$\begin{aligned} M_{\mathcal{E}} \|u - u_1\|_{H^1} \|w - \mathcal{I}^h w\|_{H^1} &\le \frac{M_{\mathcal{E}}}{m_{\mathcal{E}}} C_{\mathcal{I}} h_0^s \|\!|u - u_1|\!\| |w|_{H^{1+s}} \\ &\le \frac{M_{\mathcal{E}}}{m_{\mathcal{E}}} K_R C_{\mathcal{I}} h_0^s \|\!|u - u_1|\!\| \|u - u_1\|_{L_2}. \end{aligned} \tag{4.11}$$

For the second term of (4.10), apply (4.6) followed by (4.1) and coercivity to the interpolation error yielding

$$\begin{aligned} B\|u - u_1\|_{L_2} \|w - \mathcal{I}^h w\|_{L_2} &\le B\hat{C}_{\mathcal{I}} h_0^{1+s} \|u - u_1\|_{L_2} |w|_{H^{1+s}} \\ &\le K_R \hat{C}_{\mathcal{I}} (B h_0) h_0^s \|u - u_1\|_{L_2} \|u - u_1\|_{L_2} \\ &\le m_{\mathcal{E}}^{-1} K_R \hat{C}_{\mathcal{I}} (B h_0) h_0^s \|u - u_1\|_{L_2} \|\!|u - u_1|\!\|. \end{aligned} \tag{4.12}$$

Applying (4.11) and (4.12) to (4.10)

$$\|u - u_1\|_{L_2} \le m_{\mathcal{E}}^{-1} K_R \left(M_{\mathcal{E}} C_{\mathcal{I}} + \hat{C}_{\mathcal{I}}(B h_0) \right) h_0^s \|\!|u - u_1|\!\|. \tag{4.13}$$

□

Remark 4.2. *As the dual problem as given by* (4.3) *and* (4.4) *changes at each iteration, so may the regularity constant as given by* (4.1) *as well as the interpolation constants as given by* (4.5) *and* (4.6)*. 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.*

Remark 4.3 (Membership in H^{1+s})**.** *Depending on the regularity of the boundary $\partial\Omega$ the solution w to* (4.4) *may have less than H^2 regularity: $w \in H^2_{loc(\Omega)}$ but $w \notin H^2(\Omega)$. In particular, we may have $w \in H^{1+s}$ for some $s \in (0, 1)$. In particular, if Ω is a nonconvex polyhedral domain, then 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 $\pi/(p_j + 1) \le \omega \le \pi/p_j$ for a set of integers p_j characterizing the corners of $\partial\Omega$

$$\|w - \mathcal{I}^h w\|_{H^1} \le C h^s |w|_{1+s}$$

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* [21]*.*

Lemma 4.4 (Quasi-orthogonality)**.** *Let the problem data satisfy Assumption 2.1 and Assumption 2.3 and the mesh satisfy Assumption 2.2. Let $\mathcal{T}_1, \mathcal{T}_2 \in \mathbb{T}$ with $\mathcal{T}_2 \ge \mathcal{T}_1$. Let $u \in H^1_0(\Omega)$ the solution to* (1.2)*, $u_k \in \mathbb{V}_k$ the solution to* (2.19)*, $k = 1, 2$ and $\bar{v} \in \mathbb{V}_2$ arbitrary. Let B the constant given in Proposition 2.4. There exists a constant $C_* > 0$ depending on the problem data D and initial mesh $\mathcal{T}_0$, and a number $0 < s \le 1$ related to the angles of $\partial\Omega$, such that if the meshsize h_0 of the initial mesh satisfies $\bar{\Lambda} := B m_{\mathcal{E}}^{-1} C_* h_0^s < 1$, then*

$$|\!|\!| u - u_2 |\!|\!|^2 \le \Lambda |\!|\!| u - \bar{v} |\!|\!|^2 - |\!|\!| u_2 - \bar{v} |\!|\!|^2, \tag{4.14}$$

and in particular for $\bar{v} = u_1$

$$|\!|\!| u - u_2 |\!|\!|^2 \le \Lambda_G |\!|\!| u - u_1 |\!|\!|^2 - |\!|\!| u_2 - u_1 |\!|\!|^2, \tag{4.15}$$

where

$$\Lambda := (1 - B m_{\mathcal{E}}^{-1} C_* h_0^s)^{-1} \ \text{ and } \ \Lambda_G := (1 - B C_*^2 h_0^{2s})^{-1}$$

and C_* is the constant from Lemma 4.1.

Proof. Recombining terms

$$\begin{aligned} |\!|\!| u - u_2 |\!|\!|^2 &= a(u - \bar{v} + (\bar{v} - u_2), u - \bar{v} + (\bar{v} - u_2)) \\ &= |\!|\!| u - \bar{v} |\!|\!|^2 + |\!|\!| \bar{v} - u_2 |\!|\!|^2 + 2a(u - \bar{v}, \bar{v} - u_2) \\ &= |\!|\!| u - \bar{v} |\!|\!|^2 - |\!|\!| \bar{v} - u_2 |\!|\!|^2 + 2a(u - u_2, \bar{v} - u_2). \end{aligned} \tag{4.16}$$

By Galerkin orthogonality

$$a(u - u_2, v) + \langle b(u) - b(u_2), v \rangle = 0 \text{ for all } v \in \mathbb{V}_2. \tag{4.17}$$

Taking $v = \bar{v} - u_2$ in (4.17), by Hölder inequality and the Lipschitz assumption on b

$$\begin{aligned} 2a(u - u_2, \bar{v} - u_2) &\le 2|\langle b(u) - b(u_2), \bar{v} - u_2 \rangle| \\ &\le 2B \|u - u_2\|_{L_2} \|\bar{v} - u_2\|_{L_2}. \end{aligned} \tag{4.18}$$

In the case of (4.14) applying L_2-lifting 4.1 to the first factor on the RHS and (2.12) coercivity to the second followed by Young's inequality

$$2B\|u - u_2\|_{L_2}\|\bar{v} - u_2\|_{L_2} \leq 2Bm_{\mathcal{E}}^{-1}C_* h_0^s |\!|\!| u - u_2 |\!|\!| \, |\!|\!| \bar{v} - u_2 |\!|\!| \\ \leq Bm_{\mathcal{E}}^{-1}C_* h_0^s |\!|\!| u - u_2 |\!|\!|^2 + Bm_{\mathcal{E}}^{-1}C_* h_0^s |\!|\!| \bar{v} - u_2 |\!|\!|^2. \quad (4.19)$$

Applying (4.19) via (4.18) to (4.16)

$$(1 - Bm_{\mathcal{E}}^{-1}C_* h_0^s)|\!|\!| u - u_2 |\!|\!|^2 \leq |\!|\!| u - \bar{v} |\!|\!|^2 - (1 - Bm_{\mathcal{E}}^{-1}C_* h_0^s)|\!|\!| \bar{v} - u_2 |\!|\!|^2.$$

Assuming

$$\bar{\Lambda} \coloneqq Bm_{\mathcal{E}}^{-1}C_* h_0^s < 1 \quad (4.20)$$

we have

$$|\!|\!| u - u_2 |\!|\!|^2 \leq \Lambda |\!|\!| u - \bar{v} |\!|\!|^2 - |\!|\!| \bar{v} - u_2 |\!|\!|^2 \quad (4.21)$$

with $\Lambda = (1 - Bm_{\mathcal{E}}^{-1}C_* h_0^s)^{-1}$.

In the case of (4.15) applying L_2-lifting 4.1 to each norm on the RHS of (4.18) then applying Young's inequality

$$2B\|u - u_2\|_{L_2}\|u_1 - u_2\|_{L_2} \leq 2Bh_0^{2s}C_*^2 |\!|\!| u - u_2 |\!|\!| \, |\!|\!| u_1 - u_2 |\!|\!| \\ \leq Bh_0^{2s}C_*^2 |\!|\!| u - u_2 |\!|\!|^2 + BC_*^2 h_0^{2s} |\!|\!| u_1 - u_2 |\!|\!|^2. \quad (4.22)$$

Following the same procedure as above yields

$$|\!|\!| u - u_2 |\!|\!|^2 \leq \Lambda_G |\!|\!| u - u_1 |\!|\!|^2 - |\!|\!| u_1 - u_2 |\!|\!|^2 \quad (4.23)$$

with $\Lambda_G = (1 - BC_*^2 h_0^{2s})^{-1}$ with the weaker mesh assumption $\bar{\Lambda}_G \coloneqq BC_*^2 h_0^{2s} < 1$. $\square$

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 [16] and [2] with a clear generalization to the approximate dual sequence. Also see [6] and [18].

Lemma 4.5 (Error estimator as global upper-bound)**.** *Let the problem data satisfy Assumption 2.1 and Assumption 2.3 and the mesh satisfy Assumption 2.2. Let* $\mathcal{T}_1, \mathcal{T}_2 \in \mathbb{T}$ *with* $\mathcal{T}_2 \geq \mathcal{T}_1$. *Let* $u_1 \in \mathbb{V}_1$ *the solution to* (2.19) *and* u *the solution to* (1.2). *Let* $\hat{z}_1 \in \mathbb{V}_1$ *the solution to* (2.21)*, and* $\hat{z}$ *the solution to* (2.5)*. Then there is a global constant* C_1 *depending on the problem data* D *and initial mesh* $\mathcal{T}_0$ *with*

$$|\!|\!| u - u_1 |\!|\!| \leq C_1 \eta_1(u_1, \mathcal{T}_1) \quad (4.24)$$

and

$$|\!|\!| \hat{z} - \hat{z}_1 |\!|\!| \leq C_1 \zeta_1(\hat{z}_1, \mathcal{T}_1). \quad (4.25)$$

4.3. **Estimator Reduction.** The local Lipschitz property as in [16], analogous to the local perturbation property established in [6], is a key step in establishing estimator reduction leading to the contraction result.

Lemma 4.6 (Local Lipschitz property)**.** *Let the problem data satisfy Assumption 2.1 and Assumption 2.3 and the mesh satisfy condition (1) of Assumption 2.2. Let* B *the constant given in Proposition 2.4. Let* $\mathcal{T} \in \mathbb{T}$*. For all* $T \in \mathcal{T}$ *and for any* $v, w \in \mathbb{V}_{\mathcal{T}}$

$$\eta_{\mathcal{T}}(v, T) \leq \eta_{\mathcal{T}}(w, T) + \bar{\Lambda}_1 \eta_{\mathcal{T}}(D, T)\|v - w\|_{H^1(\omega_T)} \quad (4.26)$$

where recalling (2.18) ω_T *is the union of* T *with elements in* $\mathcal{T}$ *sharing a true-hyperface with* T*. The constant* $\bar{\Lambda}_1 > 0$ *depends on the dimension* d *and the regularity of the initial mesh* $\mathcal{T}_0$*.*

The proof follows those in [6] and [14]. The main steps are shown here.

Proof of (4.26). From (3.7)

$$\eta_{\mathcal{T}}^2(v,T) := h_T^2\|R(v)\|_{L_2(T)}^2 + h_T\|J_T(v)\|_{L_2(\partial T)}^2, \quad v \in \mathbb{V}_{\mathcal{T}}. \tag{4.27}$$

Denote $\eta_{\mathcal{T}}(v,T)$ by $\eta(v,T)$. Set $e = v - w$. Applying linearity and a generalized Taylor expansion to the definition of the residual as given by (3.1) and (3.2)

$$\begin{aligned} R(v) &= f + \mathcal{N}(w+e) \\ &= f + \nabla\cdot(A\nabla w) - b(w) + \nabla\cdot(A\nabla e) - \int_0^1 b'(w+\xi e)\,d\xi e \\ &= R(w) + \nabla\cdot(A\nabla e) - \int_0^1 b'(w+\xi e)\,d\xi e \\ &= R(w) + \mathcal{D}(e), \end{aligned}$$

where $\mathcal{D}(e) := \nabla\cdot(A\nabla e) - \int_0^1 b'(w+\xi e)\,d\xi e$. Using the generalized triangle-inequality

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

and linearity of the jump residual we have

$$\begin{aligned} \eta(v,T) &= \left(h_T^2\|R(w)+\mathcal{D}(e)\|_{L_2(T)}^2 + h_T\|J(w)+J(e)\|_{L_2(\partial T)}^2\right)^{1/2} \\ &\le \eta(w,T) + h_T\|\mathcal{D}(e)\|_{L_2(T)} + h_T^{1/2}\|J(e)\|_{L_2(\partial T)}. \end{aligned} \tag{4.28}$$

Consider the second term on the RHS. By the triangle inequality

$$\|\mathcal{D}(e)\|_{L_2(T)} \le \|\nabla\cdot(A\nabla e)\|_{L_2(T)} + \left\|\int_0^1 b'(w+\xi e)\,d\xi e\right\|_{L_2(T)}. \tag{4.29}$$

As shown in [6] and [14] the diffusion term satisfies the bound

$$\begin{aligned} \|\nabla\cdot(A\nabla e)\|_{L_2(T)} &\le \|\mathrm{div}A\cdot\nabla e\|_{L_2(T)} + \|A : D^2 e\|_{L_2(T)} \\ &\le \left(\|\mathrm{div}A\|_{L_\infty(T)} + C_I h_T^{-1}\|A\|_{L_\infty(T)}\right)\|\nabla e\|_{L_2(T)}, \end{aligned} \tag{4.30}$$

where C_I is the constant associated with an inverse inequality as in [5]. The second term in (4.29) is bounded by

$$\left\|\int_0^1 b'(w+\xi e)\,d\xi e\right\|_{L_2(T)} \le B\|e\|_{L_2(T)}. \tag{4.31}$$

As shown in [6] and [14] the jump term in (4.28) satsifies

$$\begin{aligned} \|J(e)\|_{L_2(\partial T)} &\le 2(d+1)\,C_T(\bar\gamma)^{d-1}\gamma_N^{1/2}h_T^{-1/2}\|A\|_{L_\infty(\omega_T)}\|\nabla e\|_{L_2(\omega_T)} \\ &= C_J h_T^{-1/2}\|A\|_{L_\infty(\omega_T)}\|\nabla e\|_{L_2(\omega_T)}, \end{aligned} \tag{4.32}$$

where $\bar\gamma$ and γ_N are constants of proportionality with respect to the initial mesh as given by (2.14) and (2.15) and the C_T is the constant associated with the trace theorem as in [11].

Putting together (4.28), (4.30), (4.31) and (4.32) obtain

$$\begin{aligned} \eta(v,T) &\le \eta(w,T) + h_T\left(\|\mathrm{div}A\|_{L_\infty(T)} + (C_I+C_J)h_T^{-1}\|A\|_{L_\infty(\omega_T)} + B\right)\|e\|_{H^1(\omega_T)} \\ &\le \eta(w,T) + C_{TOT}\,\eta_T(D,T)\|v-w\|_{H^1(\omega_T)}. \end{aligned} \tag{4.33}$$

□

The local perturbation property as demonstrated in Lemmas 4.6 and 5.6 leads to estimator reduction, one of the three key ingredients for contraction of the both the primal and combined quasi-errors. This result for both the primal and limiting dual problems is essentially that of [6] Corollary 2.4, [14] Theorem 3.4 and [16] Lemma 7.2. It is stated here for completeness.

Theorem 4.7 (Estimator reduction)**.** *Let the problem data satisfy Assumption 2.1 and Assumption 2.3 and the mesh satisfy conditions (1) and (2) of Assumption 2.2. Let $\mathcal{T}_1 \in \mathbb{T}$, $\mathcal{M} \subset \mathcal{T}_1$ and $\mathcal{T}_2 = REFINE(\mathcal{T}_1, \mathcal{M})$. Let*

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

with $\bar{\Lambda}_1$ from Lemma 4.6 (local Lipschitz property). Then for any $v_1 \in \mathbb{V}_1$ and $v_2 \in \mathbb{V}_2$ and $\delta > 0$

$$\eta_2^2(v_2, \mathcal{T}_2) \leq (1+\delta)\left\{\eta_1^2(v_1, \mathcal{T}_1) - \lambda \eta_1^2(v_1, \mathcal{M})\right\} + (1+\delta^{-1})\Lambda_1 \eta_0^2 |\!|\!| v_2 - v_1 |\!|\!|^2. \quad (4.34)$$

Analogously for the limiting dual problem

$$\zeta_2^2(v_2, \mathcal{T}_2) \leq (1+\delta)\left\{\zeta_1^2(v_1, \mathcal{T}_1) - \lambda \zeta_1^2(v_1, \mathcal{M})\right\} + (1+\delta^{-1})\Lambda_1 \eta_0^2 |\!|\!| v_2 - v_1 |\!|\!|^2. \quad (4.35)$$

The contraction of the primal (semilinear) problem is established in [16] and [2] based on satisfying Lemma 4.4, Lemma 4.5 and Theorem 4.7 as above. We state the result here and use it to establish our main result, Theorem 5.11.

Theorem 4.8 (Contraction of the primal problem)**.** *Let the problem data satisfy Assumption 2.1 and Assumption 2.3 and the mesh satisfy Assumption 2.2. Let u the solution to* (1.2)*. Let $\theta \in (0,1]$, and let $\{\mathcal{T}_j, \mathbb{V}_j, u_j\}_{j\geq 0}$ 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 $\mathcal{T}_0$ and marking parameter θ such that*

$$|\!|\!| u - u_{j+1} |\!|\!|^2 + \gamma_p \eta_{j+1}^2 \leq \alpha^2 \left(|\!|\!| u - u_j |\!|\!|^2 + \gamma_p \eta_j^2 \right). \quad (4.36)$$

5. Contraction and convergence of the quantity of interest

In addition to the contraction of the primal error as shown in §4, we require the analogous convergence results for the limiting dual problem: 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.

Remark 5.1. *The dual part of the combined quasi-error is written in terms of the limiting dual problem in both energy error and estimator. As such, the three convergence results listed above need only be satisfied by the limiting dual problem. As the limiting, approximate, and linearized dual problems differ only by the definition of reaction coefficient, given respectively by*

$$b'(u), \quad b'(u_j), \quad \int_0^1 b'(u_j + (u - u_j)\xi)\, d\xi,$$

it follows that the same types of estimates that hold for the limiting dual hold as well for the approximate and linearized dual sequences. This is noted in Remarks 5.3 and 5.5 with respect to the approximate dual sequence. The corresponding estimates for the linearized dual sequence are not mentioned as this sequence of problems does not naturally arise in the present convergence analysis.

To complete the analysis, we introduce the Lemma 5.9, converting between limiting and approximate estimators in order to apply the Dörfler property to a computable quantity; and Lemma 5.10, bounding the discrete error between approximate and limiting dual solutions in terms of the primal error.

We put these results together in Theorem 5.11 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 5.12.

5.1. **Limiting-dual quasi-orthogonality.**

Lemma 5.2 (Limiting-dual L_2-lifting)**.** *Let the problem data satisfy Assumption 2.1 and Assumption 2.3 and the mesh satisfy Assumption 2.2. Let $\mathcal{T}_1 \in \mathbb{T}$. Let $\hat{z} \in H_0^1(\Omega)$ the solution to* (2.5) *and $\hat{z}_1 \in \mathbb{V}_1$ the solution to* (2.21)*. Let B the constant given in Proposition 2.4. Assume for any $g \in L_2(\Omega)$ the solution y to the limiting problem: find $y \in H_0^1(\Omega)$ such that*

$$a(y, v) + \langle b'(u)y, v\rangle = g(v) \text{ for all } v \in H_0^1(\Omega) \tag{5.1}$$

belongs to $H^{1+s}(\Omega) \cap H_0^1(\Omega)$ for some $0 < s < 1$ and

$$|y|_{H^{1+s}(\Omega)} \leq \bar{K}_R \|g\|_{L_2(\Omega)}. \tag{5.2}$$

Then

$$\|\hat{z} - \hat{z}_1\| \leq \hat{C}_* h_0^s \|\hat{z} - \hat{z}_1\|. \tag{5.3}$$

Proof. The proof follows that of Lemma 4.1. As in (4.13) obtain for the limiting dual estimate (5.3)

$$\|\hat{z} - \hat{z}_1\|_{L_2} \leq m_{\mathcal{E}}^{-1} \bar{K}_R \left(M_{\mathcal{E}} \bar{C}_{\mathcal{I}} + \breve{C}_{\mathcal{I}}(B h_0)\right) h_0^s \|\hat{z} - \hat{z}_1\|. \tag{5.4}$$

□

Remark 5.3. *Under the analogous regularity assumption, Lemma 5.2 holds for the error $\hat{z}^j - \hat{z}_1^j$ in the approximate dual sequence as defined by problems* (2.4) *and* (2.20)*, respectively.*

Lemma 5.4 (Limiting-dual quasi-orthogonality)**.** *Let the problem data satisfy Assumption 2.1 and the mesh satisfy Assumption 2.2. Let $\mathcal{T}_1, \mathcal{T}_2 \in \mathbb{T}$ with $\mathcal{T}_2 \geq \mathcal{T}_1$. Let $\hat{z} \in H_0^1(\Omega)$ the solution to* (2.5) *and $\hat{z}_k \in \mathbb{V}_k$ the solution to* (2.21)*, $k = 1, 2$. Let $\bar{v} \in \mathbb{V}_2$ arbitrary. Let B the constant given in Proposition 2.4. There exists a constant $\hat{C}_* > 0$ depending on the problem data D and initial mesh $\mathcal{T}_0$, 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 $\bar{\Lambda}_* := B m_{\mathcal{E}}^{-1} \hat{C}_* < 1$, then*

$$|||\hat{z} - \hat{z}_2|||^2 \leq \hat{\Lambda} |||\hat{z} - \bar{v}|||^2 - |||\hat{z}_2 - \bar{v}|||^2 \tag{5.5}$$

and in particular for $\bar{v} = \hat{z}_1^j$ (respectively $\hat{z}_1$)

$$|||\hat{z} - \hat{z}_2|||^2 \leq \hat{\Lambda}_G |||\hat{z} - \hat{z}_1|||^2 - |||\hat{z}_2 - \hat{z}_1|||^2 \tag{5.6}$$

where

$$\hat{\Lambda} := (1 - B m_{\mathcal{E}}^{-1} \hat{C}_* h_0^s)^{-1} \text{ and } \hat{\Lambda}_G := (1 - B \hat{C}_*^2 h_0^{2s})^{-1}$$

and $\hat{C}_*$ is the constant from Lemma 5.2.

Proof. The proof follows Lemma 4.4, quasi-orthogonality in the primal problem, except in place of the inequality in (4.17) we have for the limiting dual problem

$$a(u-u_2,v)+\langle b'(u)(\hat{z}-\hat{z}_2),v\rangle = 0 \text{ for all } v\in\mathbb{V}_2, \tag{5.7}$$

yielding

$$2a(\hat{z}-\hat{z}_2,\bar{v}-\hat{z}_2)\le 2B\|\hat{z}-\hat{z}_2\|_{L_2}\|\bar{v}-\hat{z}_2\|_{L_2}, \tag{5.8}$$

as in (4.18). □

Remark 5.5. *By the same reasoning quasi-orthogonality as given by*

$$|\!|\!|\hat{z}^j-\hat{z}_2^j|\!|\!|^2\le\hat{\Lambda}|\!|\!|\hat{z}^j-\bar{v}|\!|\!|^2-|\!|\!|\hat{z}_2^j-\bar{v}|\!|\!|^2$$

and

$$|\!|\!|\hat{z}^j-\hat{z}_2^j|\!|\!|^2\le\hat{\Lambda}_G|\!|\!|\hat{z}^j-\hat{z}_1^j|\!|\!|^2-|\!|\!|\hat{z}_2^j-\hat{z}_1^j|\!|\!|^2$$

holds in the approximate dual sequence as defined by problems (2.4) *and* (2.20)*, respectively.*

5.2. **Dual sequence estimator perturbations.** The Local Lipschitz property, Lemma 4.6 (dually, 5.6) is the necessary tool to derive the estimator reduction property used to convert between estimators on different refinement levels in both the primal and limiting dual problems as in Theorem 4.7. Lemma 5.6 additionally leads to two corollaries used in the main contraction argument, Theorem 5.11 where we convert as well between estimators of the approximate and limiting problems on the same refinement level. Corollary 5.7 addresses error induced by switching between dual indicators. Then Corollary 5.8 is an immediate consequence squaring the result of Corollary 5.7 and summing over the elements. It is stated here for convenience.

Lemma 5.6 (Dual sequence local Lipschitz property)**.** *Let the problem data satisfy Assumption 2.1 and Assumption 2.3 and the mesh satisfy condition (1) of Assumption 2.2. Let $\mathcal{T}\in\mathbb{T}$. For all $T\in\mathcal{T}$ and for any $v,w\in\mathbb{V}_{\mathcal{T}}$*

$$|\zeta_{\mathcal{T},j}(v,T)-\zeta_{\mathcal{T},j}(w,T)|\le\bar{\Lambda}_1\eta_{\mathcal{T}}(D,T)\|v-w\|_{H^1(\omega_T)}. \tag{5.9}$$

In particular, for the limiting estimator

$$|\zeta_{\mathcal{T}}(v,T)-\zeta_{\mathcal{T}}(w,T)|\le\bar{\Lambda}_1\eta_{\mathcal{T}}(D,T)\|v-w\|_{H^1(\omega_T)}. \tag{5.10}$$

The constant $\bar{\Lambda}_1>0$ depends on the dimension d and the regularity of the initial mesh $\mathcal{T}_0$.

The proof follows those in [6], [14] and is nearly identical to Lemma 4.6 and is sketched here.

Proof of (5.9). From (3.8)

$$\zeta_{\mathcal{T},j}^2(v,T)\coloneqq h_T^2\|\hat{R}_j^*(v)\|_{L_2(T)}^2+h_T\|J_T(v)\|_{L_2(\partial T)}^2,\quad v\in\mathbb{V}_{\mathcal{T}}. \tag{5.11}$$

Set $e=v-w$. Applying linearity to the definition of the dual residual as given by (3.3) - (3.4)

$$\hat{R}_j^*(v)=g+\hat{\mathcal{L}}_j^*(w+e)=\hat{R}_j^*(w)+\hat{\mathcal{L}}_j^*(e).$$

By the same reasoning as (4.28)

$$\zeta_{\mathcal{T},j}(v,T)\le\zeta_{\mathcal{T},j}(w,T)+h_T\|\hat{\mathcal{L}}_j^*(e)\|_{L_2(T)}+h_T^{1/2}\|J(e)\|_{L_2(\partial T)}. \tag{5.12}$$

The term $\hat{\mathcal{L}}_j^*$ (respectively $\hat{\mathcal{L}}^*$ for the limiting dual) in (5.12) satisfies the same bound as the analogous term $\mathcal{D}$ in (4.28) of Lemma 4.6. Hence the bounds (5.9) and (5.10) hold with the same constants as in (4.26). □

Corollary 5.7 (Perturbation over approximate dual problems)**.** *Let the problem data satisfy Assumption 2.1 and Assumption 2.3 and the mesh satisfy conditions (1) and (2) of Assumption 2.2. Let $\mathcal{T} \in \mathbb{T}$, u_j the solution to* (2.19) *and u the solution to* (1.2)*. Let Θ and K_Z the constants given in Proposition 2.4. For all $T \in \mathcal{T}$ and for $v, w \in \mathbb{V}_{\mathcal{T}} \cap [z_-, z_+]$ the dual indicator on $\mathcal{T}$ satisfies*

$$|\zeta_{\mathcal{T},j}(v,T) - \zeta_{\mathcal{T},k}(w,T)| \leq \bar{\Lambda}_1 \eta_{\mathcal{T}}(D,T)\|v-w\|_{H^1(\omega_T)} + \Theta K_Z h_T \|u_j - u_k\|_{L_2(T)}. \tag{5.13}$$

In particular, for $\mathcal{T} = \mathcal{T}_1$, we have for the limiting estimator

$$|\zeta_{1,1}(v,T) - \zeta_1(w,T)| \leq \bar{\Lambda}_1 \eta_1(D,T)\|v-w\|_{H^1(\omega_T)} + \Theta K_Z h_T \|u - u_1\|_{L_2(T)} \tag{5.14}$$

and

$$|\zeta_1(w,T) - \zeta_{1,1}(v,T)| \leq \bar{\Lambda}_1 \eta_1(D,T)\|v-w\|_{H^1(\omega_T)} + \Theta K_Z h_T \|u - u_1\|_{L_2(T)}. \tag{5.15}$$

Proof. Relating dual residuals

$$\begin{aligned}
\hat{R}_j^*(w) &= g + \nabla\cdot(A\nabla w) + b'(u_k)w + (b'(u_j) - b'(u_k))\,w \\
&= \hat{R}_k^*(w) + (b'(u_j) - b'(u_k))\,w.
\end{aligned} \tag{5.16}$$

Using (5.16) in the definition of the dual indicator (3.8) and applying a generalized triangle inequality

$$\begin{aligned}
\zeta_{\mathcal{T},j}(w,T) &= \left(h_T^2\|\hat{R}_k^*(w) + (b'(u_j) - b'(u_k))w\|_{L_2(T)}^2 + h_T\|J_T(w)\|_{L_2(\partial T)}^2\right)^{1/2} \\
&\leq \left(h_T^2\|\hat{R}_k^*(w)\|_{L_2(T)}^2 + h_T\|J_T(w)\|_{L_2(\partial T)}^2\right)^{1/2} + h_T\|b'(u_j) - b'(u_k)w\|_{L_2(T)} \\
&\leq \zeta_{\mathcal{T},k}(w,T) + \Theta K_Z h_T\|u_j - u_k\|_{L_2(T)}.
\end{aligned} \tag{5.17}$$

Applying (5.9) the result of Lemma 5.6 to the estimate (5.17), obtain the result (5.13). □

Corollary 5.8 (Dual perturbation over sets)**.** *Assume the hypotheses of Corollary 5.7. Then for any subsets $\mathcal{M}_1, \mathcal{M}_2 \subseteq \mathcal{T}_1$ and arbitrary $\delta_1, \delta_2, \delta_A, \delta_B > 0$*

$$\begin{aligned}
\zeta_1^2(v,\mathcal{M}_1) \geq\ & (1+\delta_1)^{-1}(1+\delta_A)^{-1}\zeta_{1,1}^2(w,\mathcal{M}_1) \\
&- (1+\delta_1)^{-1}\delta_A^{-1}\Theta^2 K_Z^2 h_0^2\|u-u_1\|_{L_2}^2 - \delta_1^{-1}\bar{\Lambda}_1^2(d+2)\eta_0^2\|v-w\|_{H^1}^2
\end{aligned} \tag{5.18}$$

and

$$\begin{aligned}
\zeta_{1,1}^2(w,\mathcal{M}_2) \geq\ & (1+\delta_2)^{-1}(1+\delta_B)^{-1}\zeta_1^2(v,\mathcal{M}_2) \\
&- (1+\delta_2)^{-1}\delta_B^{-1}\Theta^2 K_Z^2 h_0^2\|u-u_1\|_{L_2}^2 - \delta_2^{-1}\bar{\Lambda}_1^2(d+2)\eta_0^2\|v-w\|_{H^1}^2.
\end{aligned} \tag{5.19}$$

Proof. Square equation (5.14) (respectively (5.15)) applying Young's inequality twice, then sum over element $T \in \mathcal{M} \subseteq \mathcal{T}_1$. The H^1 norm is summed over all elements $T \in \mathcal{T}_1$ counting each element $d+2$ times, the maximum number of elements in each patch ω_T. □

5.3. **Contraction of GOAFEM.** The main contraction argument Theorem 5.11 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 dual and the approximate dual problems to the primal error. Motivated by estimator reduction for the limiting dual problem as in equation (4.35)

$$\zeta_2^2(\hat z_2, \mathcal{T}_2) \leq (1+\delta)\left\{\zeta_1^2(\hat z_1, \mathcal{T}_1) - \lambda \zeta_1^2(\hat z_1, \mathcal{M})\right\} + (1+\delta^{-1})\Lambda_1 \eta_0^2 \vert\kern-0.25ex\vert\kern-0.25ex\vert \hat z_2 - \hat z_1 \vert\kern-0.25ex\vert\kern-0.25ex\vert^2 \quad (5.20)$$

the following lemma addresses the conversion between $\zeta_1^2(\hat z_1, \mathcal{M})$ and and the computable sequential estimator $\zeta_{1,1}^2(\hat z_1^1, \mathcal{M})$ necessary for marking the mesh for refinement.

Lemma 5.9 (Applying the Dörfler property to the limiting estimator)**.** *Let the problem data satisfy Assumption 2.1 and Assumption 2.3 and the mesh satisfy Assumption 2.2. Let Θ and K_Z as given by Proposition 2.4, C_* as given by Lemma 4.1 and Λ_1 as given in Lemma 4.7. Let*

u *the solution to* (1.2), u_1 *the solution to* (2.19),

$\hat z$ *the solution to* (2.5), $\hat z_1$ *the solution to* (2.21) $\hat z_1^1$ *the solution to* (2.20).

Let $\zeta_{1,1}(\hat z_1^1, \mathcal{M})$ satisfy the Dörfler property for $\mathcal{M} \subset \mathcal{T}_1$, namely

$$\zeta_{1,1}^2(\hat z_1^1, \mathcal{M}) \geq \theta^2 \zeta_{1,1}^2(\hat z_1^1, \mathcal{T}_1).$$

Then for arbitrary $\delta_1, \delta_2, \delta_A, \delta_B > 0$ there is δ_4 as given by (5.26) *such that*

$$\begin{aligned} -\zeta_1^2(\hat z_1, \mathcal{M}) \leq & -\frac{\beta\theta^2}{(1+\delta_4)}\zeta_1^2(\hat z_1, \mathcal{T}_1) - \frac{(1-\beta)\theta^2}{(1+\delta_4)} \vert\kern-0.25ex\vert\kern-0.25ex\vert \hat z - \hat z_1 \vert\kern-0.25ex\vert\kern-0.25ex\vert^2 \\ & + \left(\frac{\theta^2}{(1+\delta_A)(1+\delta_2)\delta_B} + \frac{1}{\delta_A}\right)\frac{\Theta^2 K_Z^2 C_*^2 h_0^{2(1+s)}}{(1+\delta_1)} \vert\kern-0.25ex\vert\kern-0.25ex\vert u - u_1 \vert\kern-0.25ex\vert\kern-0.25ex\vert^2 \\ & + \left(\frac{\theta^2}{(1+\delta_1)(1+\delta_A)\delta_2} + \frac{1}{\delta_1}\right)\Lambda_1 \eta_0^2(D, \mathcal{T}_0) \vert\kern-0.25ex\vert\kern-0.25ex\vert \hat z_1 - \hat z_1^1 \vert\kern-0.25ex\vert\kern-0.25ex\vert^2. \end{aligned} \quad (5.21)$$

Proof. From Corollary 5.8, L_2-lifting 4.1 and coercivity (2.12)

$$\begin{aligned} -\zeta_1^2(\hat z_1, \mathcal{M}) \leq & -(1+\delta_1)^{-1}(1+\delta_A)^{-1}\zeta_{1,1}^2(\hat z_1^1, \mathcal{M}) \\ & + (1+\delta_1)^{-1}\delta_A^{-1}\Theta^2 K_Z^2 h_0^2 \|u - u_1\|_{L_2}^2 + \delta_1^{-1}\bar\Lambda_1^2(d+2)\eta_0^2 \|\hat z_1 - \hat z_1^1\|_{H^1}^2 \\ \leq & -(1+\delta_1)^{-1}(1+\delta_A)^{-1}\zeta_{1,1}^2(\hat z_1^1, \mathcal{M}) \\ & + (1+\delta_1)^{-1}\delta_A^{-1}\Theta^2 K_Z^2 C_*^2 h_0^{2(1+s)} \vert\kern-0.25ex\vert\kern-0.25ex\vert u - u_1 \vert\kern-0.25ex\vert\kern-0.25ex\vert^2 + \delta_1^{-1}\Lambda_1 \eta_0^2 \vert\kern-0.25ex\vert\kern-0.25ex\vert \hat z_1 - \hat z_1^1 \vert\kern-0.25ex\vert\kern-0.25ex\vert^2 \end{aligned} \quad (5.22)$$

with $\Lambda_1 := \bar\Lambda_1^2(d+2)m_{\mathcal{E}}^{-2}$. The Dörfler property may be applied to the first term on the RHS of (5.22)

$$-\zeta_{1,1}^2(\hat z_1^1, \mathcal{M}) \leq -\theta^2 \zeta_{1,1}^2(\hat z_1^1). \quad (5.23)$$

Converting back to he limiting estimator by (5.19) of Corollary 5.8

$$\begin{aligned} -\zeta_{1,1}^2(\hat z_1^1) \leq & -(1+\delta_2)^{-1}(1+\delta_B)^{-1}\zeta_1^2(\hat z_1, \mathcal{M}) \\ & + (1+\delta_2)^{-1}\delta_B^{-1}\Theta^2 K_Z^2 C_*^2 h_0^{2(1+s)} \vert\kern-0.25ex\vert\kern-0.25ex\vert u - u_1 \vert\kern-0.25ex\vert\kern-0.25ex\vert^2 + \delta_2^{-1}\Lambda_1 \eta_0^2 \vert\kern-0.25ex\vert\kern-0.25ex\vert \hat z_1 - \hat z_1^1 \vert\kern-0.25ex\vert\kern-0.25ex\vert^2. \end{aligned} \quad (5.24)$$

Define δ_4 by

$$(1+\delta_4) := (1+\delta_1)(1+\delta_2)(1+\delta_A)(1+\delta_B). \quad (5.25)$$

Then by (5.22), (5.23) and (5.24)

$$\begin{aligned} -\zeta_1^2(\hat z_1, \mathcal{M}) \leq & -\theta^2(1+\delta_4)^{-1}\zeta_1^2(\hat z_1) \\ & + \left(\theta^2(1+\delta_A)^{-1}(1+\delta_2)^{-1}\delta_B^{-1} + \delta_A^{-1}\right)(1+\delta_1)^{-1}\Theta^2 K_Z^2 C_*^2 h_0^{2(1+s)} \vert\kern-0.25ex\vert\kern-0.25ex\vert u - u_1 \vert\kern-0.25ex\vert\kern-0.25ex\vert^2 \\ & + \left(\theta^2(1+\delta_1)^{-1}(1+\delta_A)^{-1}\delta_2^{-1} + \delta_1^{-1}\right)\Lambda_1 \eta_0^2 \vert\kern-0.25ex\vert\kern-0.25ex\vert \hat z_1 - \hat z_1^1 \vert\kern-0.25ex\vert\kern-0.25ex\vert^2. \end{aligned} \quad (5.26)$$

Finally, for $\beta \in (0,1)$ split the first term in (5.26) into two pieces, applying the upper-bound estimate from Lemma 4.5 to the second piece yielding

$$\begin{aligned}-\zeta_1^2(\hat{z}_1, \mathcal{M}) \leq &-\beta\theta^2(1+\delta_4)^{-1}\zeta_1^2(\hat{z}_1) - (1-\beta)\theta^2(1+\delta_4)^{-1}C_1^{-2}|\!|\!|\hat{z} - \hat{z}_1|\!|\!|^2 \\ &+ \left(\theta^2(1+\delta_A)^{-1}(1+\delta_2)^{-1}\delta_B^{-1} + \delta_A^{-1}\right)(1+\delta_1)^{-1}\Theta^2 K_Z^2 C_*^2 h_0^{2(1+s)}|\!|\!|u - u_1|\!|\!|^2 \\ &+ \left(\theta^2(1+\delta_1)^{-1}(1+\delta_A)^{-1}\delta_2^{-1} + \delta_1^{-1}\right)\Lambda_1\eta_0^2|\!|\!|\hat{z}_1 - \hat{z}_1^1|\!|\!|^2. \end{aligned} \tag{5.27}$$

□

Lemma 5.10 (Bounding the error in the discrete problem)**.** *Let the problem data satisfy Assumption 2.1 and Assumption 2.3 and the mesh satisfy Assumption 2.2. Let* Θ *and* K_Z *the constants given in Proposition 2.4 and* C_* *and* $\hat{C}_*$ *the constants given by Lemmas 4.1 and 5.2, respectively. Let*

u *the solution to* (1.2), $\quad$ u_1 *the solution to* (2.19),

$\hat{z}_1$ *the solution to* (2.21), $\quad$ $\hat{z}_1^1$ *the solution to* (2.20).

Then

$$|\!|\!|\hat{z}_1 - \hat{z}_1^1|\!|\!| \leq \Theta K_Z C_* \hat{C}_* h_0^{2s}|\!|\!|u - u_1|\!|\!|. \tag{5.28}$$

Proof. Recall that

$$\hat{z}_1 \text{ solves } a(\hat{z}_1, v) + \langle b'(u)\hat{z}_1, v\rangle = g(v), \quad \text{for all } v \in \mathbb{V}_1 \tag{5.29}$$

$$\hat{z}_1^1 \text{ solves } a(\hat{z}_1^1, v) + \langle b'(u_1)\hat{z}_1^1, v\rangle = g(v), \quad \text{for all } v \in \mathbb{V}_1. \tag{5.30}$$

Subtracting (5.30) from (5.29) and rearranging terms

$$a(\hat{z}_1 - \hat{z}_1^1, v) + \langle (b'(u) - b'(u_1))\hat{z}_1, v\rangle = \langle b'(u_1)(\hat{z}_1^1 - \hat{z}_1), v\rangle,\ v \in \mathbb{V}_1. \tag{5.31}$$

In particular, for $v = \hat{z}_1 - \hat{z}_1^1 \in \mathbb{V}_1$ equation (5.31) yields

$$\begin{aligned}|\!|\!|\hat{z}_1 - \hat{z}_1^1|\!|\!|^2 &= -\langle (b'(u) - b'(u_1))\hat{z}_1, \hat{z}_1 - \hat{z}_1^1\rangle - \langle b'(u_1)(\hat{z}_1 - \hat{z}_1^1), \hat{z}_1 - \hat{z}_1^1\rangle \\ &\leq -\langle (b'(u) - b'(u_1))\hat{z}_1, \hat{z}_1 - \hat{z}_1^1\rangle \end{aligned} \tag{5.32}$$

where the last line in (5.32) follows from the assumption that b is an increasing function hence $\langle b'(u_1)(\hat{z}_1 - \hat{z}_1^1), \hat{z}_1 - \hat{z}_1^1\rangle \geq 0$. Then applying the Lipschitz property of b', the *a priori* bound on the dual solution $\hat{z}_1$ and both primal and dual L_2 lifting we have from (5.32)

$$\begin{aligned}|\!|\!|\hat{z}_1 - \hat{z}_1^1|\!|\!|^2 &\leq \Theta K_Z \|u - u_1\|_{L_2}\|\hat{z}_1 - \hat{z}_1^1\|_{L_2} \\ &\leq \Theta K_Z C_* \hat{C}_* h_0^{2s}|\!|\!|u - u_1|\!|\!|\,|\!|\!|\hat{z}_1 - \hat{z}_1^1|\!|\!| \end{aligned} \tag{5.33}$$

from which the result follows. □

The contraction of the combined quasi-error is driven by the dual-sequence estimator reduction and quasi-orthogonality estimates. As the former is coupled to the primal error, the end result is a reduction in a linear combination of the energy errors in primal and limiting dual problems and error estimators of the primal problem and approximate dual sequence. As seen by the bound on the error in the goal function Theorem 5.12, the contraction of the combined quasi-error determines the contraction in the error of the quantity of interest.

Theorem 5.11 (Contraction of the combined quasi-error)**.** *Let the problem data satisfy Assumption 2.1 and Assumption 2.3 and the mesh satisfy Assumption 2.2. Let*

u *the solution to* (1.2), $\quad$ u_j *the solution to* (2.19),

$\hat{z}$ *the solution to* (2.5), $\quad$ $\hat{z}_j$ *the solution to* (2.21).

Let $\theta \in (0,1]$, and let $\{\mathcal{T}_j, \mathbb{V}_j\}_{j\geq 0}$ be the sequence of meshes and finite element spaces produced by GOAFEM. Let $\gamma_p > 0$ as given by Theorem 4.8. Then there exist constants $\gamma > 0, \pi > 0$ and $0 < \alpha_D < 1$, depending on the initial mesh $\mathcal{T}_0$ and marking parameter θ such that

$$\begin{aligned}&|||\hat{z} - \hat{z}_2|||^2 + \gamma\zeta_2^2(\hat{z}_2) + \pi|||u - u_2|||^2 + \pi\gamma_p\eta_2^2(u_2)\\ &\leq \alpha_D^2\left(|||\hat{z} - \hat{z}_1|||^2 + \gamma\zeta_1^2(\hat{z}_1) + \pi|||u - u_1|||^2 + \pi\gamma_p\eta_1^2(u_1)\right).\end{aligned} \tag{5.34}$$

Proof. Let

$$\eta_0 = \eta_0(D, \mathcal{T}_0) \text{ and } \zeta_k(\hat{z}_k) = \zeta_k(\hat{z}_k, \mathcal{T}_k),\ k = 1, 2.$$

Start with estimator reduction for the limiting dual problem as in equation (4.35). For arbitrary $\delta > 0$

$$\zeta_2^2(\hat{z}_2) \leq (1+\delta)\left\{\zeta_1^2(\hat{z}_1) - \lambda\zeta_1^2(\hat{z}_1, \mathcal{M})\right\} + (1+\delta^{-1})\Lambda_1\eta_0^2|||\hat{z}_2 - \hat{z}_1|||^2. \tag{5.35}$$

Recall the quasi-orthogonality estimate in the limiting dual problem from Lemma 5.4

$$|||\hat{z} - \hat{z}_2|||^2 \leq \hat{\Lambda}_G|||\hat{z} - \hat{z}_1|||^2 - |||\hat{z}_2 - \hat{z}_1|||^2. \tag{5.36}$$

Adding (5.36) to a positive multiple γ (to be determined) of (5.35) and applying the results of Lemmas 5.9 and 5.10 obtain

$$\begin{aligned}|||\hat{z} - \hat{z}_2|||^2 + \gamma\zeta_2^2(\hat{z}_2) \leq A|||\hat{z} - \hat{z}_1|||^2 + \gamma B\zeta_1^2(\hat{z}_1) + D|||u - u_1|||^2&\\ + \left(\gamma(1+\delta^{-1})\Lambda_1\eta_0^2 - 1\right)|||\hat{z}_2 - \hat{z}_1|||^2.&\end{aligned} \tag{5.37}$$

Set γ to eliminate the last term in (5.37)

$$\gamma := (1+\delta^{-1})^{-1}\Lambda_1^{-1}\eta_0^{-2}. \tag{5.38}$$

Then the coefficients A and B of (5.36) are given by

$$A = \hat{\Lambda}_G - (1-\beta)\lambda\theta^2\delta(1+\delta_4)^{-1}C_1^{-2}\Lambda_1^{-1}\eta_0^{-2} \tag{5.39}$$

$$B = (1+\delta)(1 - \beta\lambda\theta^2(1+\delta_4)^{-1}) \tag{5.40}$$

where as given in (5.25)

$$(1+\delta_4) := (1+\delta_1)(1+\delta_2)(1+\delta_A)(1+\delta_B).$$

For the coefficients as defined by (5.39) and (5.40), requiring $A < 1$ and $B < 1$ yields the inequality

$$\frac{\delta}{1+\delta}\frac{1+\delta_4}{\lambda\theta^2} < \beta < 1 - \frac{(\hat{\Lambda}_G - 1)\Lambda_C}{\delta}\frac{1+\delta_4}{\lambda\theta^2}. \tag{5.41}$$

To demonstrate that parameters $\delta, \delta_4 > 0$ may be chosen to satisfy (5.41) with $\beta \in (0,1)$ set

$$\delta_4 = \delta = b\lambda\theta^2 \text{ for some } b < 1. \tag{5.42}$$

Require the mesh condition

$$(\hat{\Lambda}_G - 1)\Lambda_C = a\lambda\theta^2 \text{ for some } a < 1 \text{ with } \Lambda_C := C_1^2\Lambda_1\eta_0^2 \tag{5.43}$$

for a given $\theta \in (0,1)$. Then using (5.42) and (5.43) in (5.41) yields

$$b < \beta < 1 - a\left(1 + \frac{1}{b\lambda\theta^2}\right) \tag{5.44}$$

which may be satisfied with $\beta \in (0,1)$ for n sufficiently small. The condition (5.43) with a as required by (5.44) is feasible as the the dual quasi-orthogonality constant $\hat{\Lambda}_G$ may be driven arbitrarily close to unity by a sufficiently fine initial mesh.

Consider the coefficient D of (5.37). As we have conditions on the combined parameter δ_4, assume $\delta_1 = \delta_2 = \delta_A = \delta_B =: \delta_C$. Then

$$D = \delta\lambda\Theta^2 K_Z^2 C_*^2 h_0^{2s} \left(\frac{\theta^2 + (1+\delta_C)^2}{(1+\delta_C)^2\delta_C}\right)\left(\frac{h_0^2}{\Lambda_1\eta_0^2} + \hat{C}_*^2 h_0^{2s}\right). \tag{5.45}$$

To control the primal error term with the coefficient D as given by (5.45), add a positive multiple π (to be determined) of the primal contraction result (4.36) of Theorem 4.8 to (5.39) yieding

$$\begin{aligned}
&|||\hat{z} - \hat{z}_2|||^2 + \gamma\zeta_2^2(\hat{z}_2) + \pi|||u - u_2|||^2 + \pi\gamma_p\eta_2^2(u_2) \\
&\leq A|||\hat{z} - \hat{z}_1|||^2 + \gamma B\zeta_1^2(\hat{z}_1) + (D + \alpha^2\pi)|||u - u_1|||^2 + \alpha^2\pi\gamma_P\eta_1^2(u_1).
\end{aligned} \tag{5.46}$$

Choose π to ensure $D + \alpha^2\pi < \pi$

$$\pi > \frac{D}{1-\alpha^2} \tag{5.47}$$

and set

$$\alpha_D := \max\left\{A, B, \frac{D+\alpha^2\pi}{\pi}, \alpha^2\right\} < 1. \tag{5.48}$$

Then the combined quasi-error satisfies the contraction property

$$\begin{aligned}
&|||\hat{z} - \hat{z}_2|||^2 + \gamma\zeta_2^2(\hat{z}_2) + \pi|||u - u_2|||^2 + \pi\gamma_p\eta_2^2(u_2) \\
&\leq \alpha_D^2\left(|||\hat{z} - \hat{z}_1|||^2 + \gamma\zeta_1^2(\hat{z}_1) + \pi|||u - u_1|||^2 + \pi\gamma_p\eta_1^2(u_1)\right).
\end{aligned} \tag{5.49}$$

□

Corollary 5.12 (Bounding the error in the goal function)**.** *Let the problem data satisfy Assumption 2.1 and Assumption 2.3 and the mesh satisfy Assumption 2.2. Let B, Θ and K_Z the constants given in Proposition 2.4 and C_* and $\hat{C}_*$ the constants given by Lemmas 4.1 and 5.2, respectively. Let $\alpha_D \in (0,1)$ as given by Theorem 5.11. Let*

u *the solution to* (1.2),	u_j *the solution to* (2.19),
$\hat{z}$ *the solution to* (2.5),	$\hat{z}_j$ *the solution to* (2.21).

Then the error in the goal function is controlled by a constant multiple of the square of the combined quasi-error, and

$$|g(u) - g(u_j)| \leq C\bar{Q}_j^2(u_j, \hat{z}_j) \leq \alpha_D^{2j} C\bar{Q}_0^2(u_0, \hat{z}_0). \tag{5.50}$$

Proof. Choosing the test function $v = u - u_j$ in (2.5), and by linearity and Galerkin orthogonality for the primal problem

$$\begin{aligned}
g(u) - g(u_j) &= a(\hat{z}, u) + \langle b'(u)\hat{z}, u\rangle - a(\hat{z}, u_j) - \langle b'(u)\hat{z}, u_j\rangle \\
&= a(u - u_j, \hat{z}) + \langle b'(u)(u - u_j), \hat{z}\rangle \\
&= a(u - u_j, \hat{z}) + \langle \mathcal{B}_j(u - u_j), \hat{z}\rangle + \langle (b'(u) - \mathcal{B}_j)(u - u_j), \hat{z}\rangle \\
&= a(u - u_j, \hat{z} - \hat{z}_j) + \langle b(u) - b(u_j), \hat{z} - \hat{z}_j\rangle + \langle (b'(u) - \mathcal{B}_j)(u - u_j), \hat{z}\rangle.
\end{aligned} \tag{5.51}$$

The third term in the last line of (5.51) represents the error induced by switching from the limiting to the linearized dual problem as required to make use of property (2.2). This term may be bounded in terms of the constants and L_∞ estimates in Proposition 2.4 and

$$\|b'(u) - \mathcal{B}_j\| = \left\|\int_0^1 b'(u) - b'\left(u_j + \xi(u - u_j)\right) d\xi\right\| \leq \frac{\Theta}{2}\|u - u_j\|,$$

yielding

$$\begin{aligned}\langle (b'(u) - \mathcal{B}_j)(u - u_j), \hat{z}\rangle &\leq K_Z \|b'(u) - \mathcal{B}_j\|_{L_2}\|u - u_j\|_{L_2} \\ &\leq \frac{1}{2}\Theta K_Z \|u - u_j\|^2_{L_2}. \end{aligned} \tag{5.52}$$

Then by (5.51), (5.52), the Cauchy-Schwarz inequality and L_2-lifting as in Lemmas 4.1 and 5.2

$$\begin{aligned}|g(u) - g(u_j)| &\leq |||u - u_j||| \, |||\hat{z} - \hat{z}_j||| + B\|u - u_j\|_{L_2}\|\hat{z} - \hat{z}_j\|_{L_2} \\ &\quad + \frac{1}{2}\Theta K_Z \|u - u_j\|^2_{L_2} \\ &\leq (1 + BC_*\hat{C}_* h_0^{2s}) |||u - u_j||| \, |||\hat{z} - \hat{z}_j||| + \frac{1}{2}\Theta K_Z C_*^2 h_0^{2s} |||u - u_j|||^2 \\ &\leq \frac{1}{2}\left(1 + (\Theta K_Z C_* + B\hat{C}_*)C_* h_0^{2s}\right) |||u - u_j|||^2 \\ &\quad + \frac{1}{2}(1 + BC_*\hat{C}_* h_0^{2s}) |||\hat{z} - \hat{z}_j|||^2. \end{aligned} \tag{5.53}$$

Comparing (5.53) to the (5.49), the error in the goal function is bounded below a constant multiple of the combined quasi-error

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

which is shown to contract at each iteration of the algorithm, from which (5.50) follows. □

6. 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 included 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. 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.

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

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

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

[8] M. G. L. D. Estep 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.

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

[12] T. Gilbarg. *Elliptic partial differential equations of second order*. Springer-Verlag, 1977.

[13] M. Holst. Applications of domain decomposition and partition of unity methods in physics and geometry, 2003.

[14] M. Holst and S. Pollock. Convergence of goal oriented methods for nonsymmetric problems, 2011.

[15] M. Holst, R. Szypowski, and Y. Zhu. Two-grid methods for semilinear interface problems, 2012.

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

[17] S. Kesavan. *Topics in Functional Analysis and Applications*. John Wiley and Sons, Inc., New York, NY, 1989.

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

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

[22] M. Struwe. *Variational Methods*. Springer-Verlag, Berlin, Germany, 3 edition, 2000.

E-mail address: `mholst@math.ucsd.edu`

E-mail address: `snpolloc@math.ucsd.edu`

E-mail address: `zhu@math.ucsd.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA SAN DIEGO, LA JOLLA CA 92093