LOCAL ERROR ESTIMATES AND ADAPTIVE REFINEMENT FOR FIRST-ORDER SYSTEM LEAST SQUARES (FOSLS)[†]

MARKUS BERNDT[‡§], THOMAS A. MANTEUFFEL[‡¶], AND STEPHEN F. MCCORMICK[‡‖]

Abstract. We establish an a-posteriori error estimate, with corresponding bounds, that is valid for any FOSLS L^2-minimization problem. Such estimates follow almost immediately from the FOSLS formulation, but they are usually difficult to establish for other methodologies. We present some numerical examples to support our theoretical results. We also establish a local a-priori lower error bound that is useful for indicating when refinement is necessary and for determining the initial grid. Finally, we obtain a sharp theoretical error estimate under certain assumptions on the refinement region and show how this provides the basis for an effective refinement strategy. The local a-priori lower error bound and the sharp theoretical error estimate both appear to be unique to the least-squares approach.

Key words. adaptive mesh refinement, a-posteriori error estimates, first-order system least-squares

AMS subject classifications. 65N15, 65N30, 65N50

1. Introduction. Research has recently intensified in the field of first-order system least-squares methods (cf. [5], [7], [8], [6], [3], [4] and [11]). Most of the numerical examples presented in these papers are based on uniform grid implementations. However, in [11], the contribution that an element makes to the total value of the FOSLS functional is used as a local a-posteriori error estimate in a refinement process. This measure had been suggested earlier by a number of authors (eg. [12] and [9]), but the reasoning for such an a-posteriori error estimate has been heuristic. The purpose of the present paper is to put this methodology on a theoretical footing.

In Section 2, we review the FOSLS methodology to introduce some notation and motivate our later results. In Section 3, we establish the theory for such a local a-posteriori error estimate using only results from standard FOSLS theory in a general setting. Our results imply that any FOSLS L^2-functional can be used as a local a-posteriori estimate for the error in the norm that it induces. Such estimates follow almost immediately from the FOSLS formulation, but they can be difficult to establish for other methodologies. We support this theory and these claims with a numerical example. In Section 4, we show how the FOSLS functional can also be used to obtain a local a-priori lower error bound. This a-priori bound, which appears to be unique to the least-squares approach, is useful for indicating when refinement is necessary and determining the initial grid. In Section 5, we develop another unique tool: a sharp theoretical error estimate. We prove this estimate to be a good indicator of the local error under certain assumptions on the refinement region, and we show how it provides the basis for an effective refinement strategy.

2. Notation. In this section, we review the FOSLS methodology in a general setting and introduce notation used in Section 3.

† This work was sponsored by the National Science Foundation under grant number DMS-9312752, and the Department of Energy under grant number DE-FG03-94ER25217

‡ Department of Applied Mathematics, Campus Box 526, University of Colorado at Boulder, Boulder, CO 80303-0526

§ Email: berndt@boulder.colorado.edu

¶ Email: tmanteuf@boulder.colorado.edu

‖ Email: stevem@boulder.colorado.edu

2.1. The FOSLS Methodology. Here we review the FOSLS methodology. See [5] and [7] for more detail.

We start with a (typically second-order) partial differential equation, or system of partial differential equations:

$$\mathcal{L}\mathbf{w} = \mathbf{f} \text{ in } \Omega, \tag{1}$$

together with appropriate boundary conditions. The FOSLS methodology yields a system of first-order partial differential equations,

$$\mathcal{L}_i\mathbf{u} = f_i, \ i = 1 \ldots M, \tag{2}$$

which is equivalent to original problem (1). The resulting FOSLS L^2-functional is the scaled sum of L^2-norms of the residuals of system (2):

$$\mathcal{G}(\mathbf{u};\mathbf{f}) = \sum_{i=1}^{M} a_i \|\mathcal{L}_i\mathbf{u} - f_i\|_0^2, \ a_i > 0. \tag{3}$$

(Here we consider the L^2 case for definiteness, although the L^2 norm $\|\cdot\|_0$ can be replaced by any other suitable norm.) The FOSLS minimization problem is

$$\mathbf{u} = \arg\min_{\mathbf{v}\in W} \mathcal{G}(\mathbf{v};\mathbf{f}), \tag{4}$$

where W is an appropriate Hilbert space, often a product of H^1 spaces. The weak form of this minimization problem is to find $\mathbf{u} \in W$ such that

$$\mathcal{F}(\mathbf{u};\mathbf{v}) = (f, \mathbf{v})_0, \quad \forall \mathbf{v} \in W, \tag{5}$$

where

$$\mathcal{F}(\mathbf{u};\mathbf{v}) = \sum_{i=1}^{M} a_i (\mathcal{L}_i\mathbf{u}, \mathcal{L}_i\mathbf{v})_0. \tag{6}$$

The next step in the FOSLS methodology is to establish continuity and coercivity bounds for bilinear form (6) in some suitable norm $|\!|\!|\cdot|\!|\!|$. In many cases, this norm is a properly scaled sum of H^1-norms of components of $\mathbf{u}$. Continuity is

$$\mathcal{F}(\mathbf{u};\mathbf{v}) \le \bar{c}\, |\!|\!|\mathbf{u}|\!|\!|\, |\!|\!|\mathbf{v}|\!|\!| \tag{7a}$$

and coercivity is

$$\mathcal{F}(\mathbf{u};\mathbf{u}) \ge \underline{c}\, |\!|\!|\mathbf{u}|\!|\!|^2. \tag{7b}$$

Among other attributes, these bounds imply well-posedness of FOSLS minimization problem (4).

2.2. The Local FOSLS Functional. The FOSLS functional is a sum of integrals and, hence, can be evaluated over any subdomain A of Ω. We call

$$\mathcal{G}_A(\mathbf{u};\mathbf{f}) = \sum_{i=1}^{M} a_i \|\mathcal{L}_i\mathbf{u} - f_i\|_{0,A}^2 \tag{8}$$

the local FOSLS functional. It follows that, for any tessellation $\mathcal{T}$ of Ω, the FOSLS functional can be evaluated as the sum of all local functionals over the elements in the tessellation:

$$\mathcal{G}(\mathbf{u};\mathbf{f}) = \sum_{\tau\in\mathcal{T}} \mathcal{G}_\tau(\mathbf{u};\mathbf{f}). \tag{9}$$

3. A-Posteriori Error Estimates. In the literature, it has been suggested that the value of the local least-squares functional on an element in a given tessellation can be used as an a-posteriori error estimate (cf. [12], [9], [11]). From standard FOSLS theory, it is usually easy to establish bounds for such an a-posteriori error estimate.

For a given $\mathbf{u}^h$ in some finite-dimensional space $W^h \subset W$, define

$$\eta_\tau := \sqrt{\mathcal{G}_\tau(\mathbf{u}^h;\mathbf{f})} \tag{10}$$

for any $\tau \in \mathcal{T}$, where $\mathcal{T}$ is any tessellation of the domain Ω (not necessarily associated with W^h). Now, coercivity bound (7b) implies

$$\mathcal{G}(\mathbf{u}^h;\mathbf{f}) = \mathcal{G}(\mathbf{u}^h - \mathbf{u};\mathbf{0}) = \mathcal{F}(\mathbf{u}^h - \mathbf{u};\mathbf{u}^h - \mathbf{u}) \geq \underline{c}\, |\!|\!|\mathbf{u}^h - \mathbf{u}|\!|\!|^2 \tag{11}$$

and, from (9) and the definition of η_τ, we get

$$\mathcal{G}(\mathbf{u}^h;\mathbf{f}) = \sum_{\tau\in\mathcal{T}} \mathcal{G}_\tau(\mathbf{u}^h;\mathbf{f}) = \sum_{\tau\in\mathcal{T}} \eta_\tau^2. \tag{12}$$

Thus, (11) and (12) imply

$$|\!|\!|\mathbf{u}^h - \mathbf{u}|\!|\!|^2 \leq \frac{1}{\underline{c}} \sum_{\tau\in\mathcal{T}} \eta_\tau^2. \tag{13}$$

In the literature, an inequality of type (13) is called a reliability bound (cf. [13]). If all of the local error estimates η_τ are small, then the error is also small.

For FOSLS formulations, it is typically straightforward to establish bound (7a), using only the triangle, Cauchy-Schwarz, and ε inequalities. Most importantly, we will show in the next subsection by example that a proof analogous to that for bound (7a) generally can be used to establish a similar bound on the local functional. We get

$$\mathcal{G}_A(\mathbf{u}^h;\mathbf{f}) = \mathcal{F}_A(\mathbf{u}^h - \mathbf{u};\mathbf{u}^h - \mathbf{u}) \leq \bar{c}|\!|\!|\mathbf{u}^h - \mathbf{u}|\!|\!|_A^2 \tag{14}$$

for any subdomain A of Ω, which implies

$$|\!|\!|\mathbf{u}^h - \mathbf{u}|\!|\!|_\tau^2 \geq \frac{1}{\bar{c}}\, \eta_\tau^2 \tag{15}$$

for any τ in $\mathcal{T}$.

For these estimates, we made no assumption about how the approximation $\mathbf{u}^h \in W^h$ was obtained. For standard Galerkin formulations, bounds like (13) and (15) for an a-posteriori error estimate usually depend on the fact that $\mathbf{u}^h$ comes from a discrete finite element solution, and they can be very tedious to derive (cf. [13] for a number of examples). In contrast, the choice of (10) as an a-posteriori error estimate for a FOSLS formulation is natural, since bounds (13) and (15) follow immediately from FOSLS theory. Coercivity bound (7b) yields bound (13) as we showed above. A proof analogous to that for continuity bound (7a) yields (15), as we now illustrate by two examples.

3.1. Examples. Here we include examples that illustrate how inequality (15) can be established in general. First consider the div-curl functional in 3D:

$$\mathcal{G}(\mathbf{u};f) = \|\nabla\cdot\mathbf{u} + f\|_{0,\Omega}^2 + \|\nabla\times\mathbf{u}\|_{0,\Omega}^2. \tag{16}$$

To understand our notation, see [7]. From vector calculus and the triangle, Cauchy-Schwarz, and ε inequalities, we obtain

$$\|\nabla \cdot \mathbf{e}\|_{0,A}^2 \leq 3\, \|\nabla \mathbf{e}\|_{0,A}^2, \tag{17a}$$

$$\|\nabla \times \mathbf{e}\|_{0,A}^2 \leq 3\, \|\nabla \mathbf{e}\|_{0,A}^2. \tag{17b}$$

Here, $\mathbf{e} \in (H^1(\Omega))^3$ and A is any subdomain of Ω. These two inequalities and the definition of $\mathcal{G}$ imply

$$\begin{aligned} \eta_\tau^2 &= \mathcal{G}_\tau(\mathbf{u}^h; f) \\ &= \|\nabla \cdot (\mathbf{u}^h - \mathbf{u})\|_{0,\tau}^2 + \|\nabla \times (\mathbf{u}^h - \mathbf{u})\|_{0,\tau}^2 \\ &\leq 6\, \|\nabla(\mathbf{u}^h - \mathbf{u})\|_{0,\tau}^2, \end{aligned} \tag{18}$$

which immediately implies (15) with $\bar{c} = 6$.

Next, as a more realistic example, consider the Stokes functional discussed in [8]:

$$\begin{aligned} \mathcal{G}(\underline{U}, \mathbf{u}, p; \mathbf{f}, g) \;=\; & \|\mathbf{f} + \nu(\nabla \cdot \underline{U})^t - \nabla p\|_0^2 + \nu^2 \|\underline{U} - \nabla \mathbf{u}^t\|_0^2 + \nu^2 \|\nabla \times \underline{U}\|_0^2 \\ & + \nu^2 \|\nabla \cdot \mathbf{u} - g\|_0^2 + \nu^2 \|\nabla(tr\underline{U}) - \nabla g\|_0^2. \end{aligned} \tag{19}$$

Similar to the div-curl example, we can establish a bound of type (15) using only the triangle, Cauchy-Schwarz, and ε inequalities:

$$\frac{1}{12}\, \eta_\tau^2 \leq \nu^2\, \|\underline{U}^h - \underline{U}\|_{1,\tau}^2 + \nu^2\, \|\mathbf{u}^h - \mathbf{u}\|_{1,\tau}^2 + \|p^h - p\|_{1,\tau}^2. \tag{20}$$

These two examples are typical and show that bounds (13) and (15) follow naturally from standard FOSLS theory.

3.2. Numerical Results. Here we present a simple heuristic refinement strategy that incorporates the a-posteriori error estimate (3) to illustrate its practical value. An optimal strategy determines a refinement region $R \subset \Omega$ that minimizes the ratio

$$\frac{\text{work to solve the new discrete problem}}{\text{gain in accuracy}}. \tag{21}$$

Of course, we must somehow approximate both work and gain. For a FOSLS implementation, the solver of choice is a multilevel algorithm. Hence, a reasonable approximation to work is a linear function of the number of vertices in the finite element mesh, n_{old}, and the number of vertices that will be added due to refinement, n_{new}. The new vertices are added to the hierarchy of levels as the finest level. Hence, the work induced by these points (relaxation and grid transfers) is proportional to n_{new}. The work induced by the old vertices is proportional to n_{old}, since a multigrid cycle is used to solve the coarse level problem. Thus, the following approximation for work is used:

$$\text{work to solve the new discrete problem} \approx a n_{old} + b n_{new}, \tag{22}$$

with suitable constants a and b. The gain in accuracy can be measured by calculating the ratio

$$\text{gain in accuracy} = \frac{\mathcal{G}(\mathbf{u}_{old}^h; \mathbf{f})}{\mathcal{G}(\mathbf{u}_{new}^h; \mathbf{f})}. \tag{23}$$

We cannot calculate $\mathcal{G}(\mathbf{u}^h_{new};\mathbf{f})$ before refining the mesh, which we assume is accomplished by halving the mesh size: mesh size h is reduced to $\frac{h}{2}$. For a FOSLS discretization of order h, we typically get bounds like $\|e\| \leq h\ \|\nabla u\|$, so cutting the mesh size h in half tends to reduce the functional by a factor of 4. Thus, we use the approximation

$$\mathcal{G}(\mathbf{u}^h_{new};\mathbf{f}) \approx \frac{1}{4}\mathcal{G}_R(\mathbf{u}^h_{old};\mathbf{f}) + \mathcal{G}_{R^c}(\mathbf{u}^h_{old};\mathbf{f}). \tag{24}$$

Our refinement strategy is as follows.

1. Find the maximum of the a-posteriori error estimates over all $\tau \in \mathcal{T}$: $\eta_{\mathcal{T},\max}$.
2. Partition $\mathcal{T}$ into contour sets $C_i = \{\tau | \eta^2_\tau \in (\frac{i-1}{N}\eta^2_{\mathcal{T},\max}, \frac{i}{N}\eta^2_{\mathcal{T},\max}]\}$, $i = 1, 2, \ldots, N$.
3. Calculate $w_i = \text{work/gain}$, for $i = 1, 2, \ldots, N$.
4. Find $i \in \{1, 2, \ldots, N\}$ for which w_i is minimal.
5. Refine contour sets $C_i \ldots C_N$.

We applied this refinement strategy to the FOSLS minimization problem involving functional (16) on the unit square with boundary conditions $\tau \cdot \mathbf{u} = 0$ (which corresponds with $\mathbf{u} = \nabla p$ to a Poisson equation with homogeneous Dirichlet boundary conditions). In this example, we chose $f = \chi_{B_{0.1}(\frac{3}{4},\frac{3}{4})}$, the characteristic function of a ball of radius 0.1 centered at the point $(\frac{3}{4}, \frac{3}{4})$. We used continuous piecewise linear elements that conformed to the boundary conditions. Table 1 shows the value

iteration	# unknowns	$\mathcal{G}(\mathbf{u}; f)$ (loc. ref.)	$\mathcal{G}(\mathbf{u}; f)$ (glob. ref.)
0	62	8.05(-3)	8.05(-3)
1	94	7.38(-3)	7.51(-3)
2	214	3.99(-3)	4.68(-3)
3	468	2.14(-3)	3.59(-3)
4	866	1.36(-3)	2.75(-3)
5	2580	6.56(-4)	1.64(-3)
6	4705	3.80(-4)	1.15(-3)

TABLE 1
Value of FOSLS functional: local versus global refinement

of the global FOSLS functional after solving the minimization problem on the locally refined finite element mesh. As a reference, we also show the value of the global FOSLS functional after solving on a uniform finite element mesh with approximately the same number of vertices as the corresponding refined mesh. Our results show that the a-posteriori error estimate (10) is a good indicator for the necessity of refinement.

4. A Local A-Priori Lower Error Bound. Here we show how the local FOSLS functional can be used to obtain a local a-priori error estimate.

For a given element τ, define the estimate

$$\bar{\eta}_\tau := \sqrt{\min_{\mathbf{v}^h \in W^h} \mathcal{G}_\tau(\mathbf{v}^h;\mathbf{f})}. \tag{25}$$

Then, for any tessellation $\mathcal{T}$ that contains τ, we have

$$\bar{\eta}^2_\tau \leq \mathcal{G}_\tau(\mathbf{v}^h;\mathbf{f}), \quad \forall \mathbf{v}^h \in W^h_{\mathcal{T}}. \tag{26}$$

Using (15), this implies that

$$\bar{\eta}_\tau^2 \le \eta_\tau^2 \le \bar{c}\, \|\mathbf{u}^h - \mathbf{u}\|_\tau^2 \tag{27}$$

for any tessellation $\mathcal{T}$ with its associated discrete space $W^h_{\mathcal{T}}$.

A-priori error estimate $\bar{\eta}_\tau$ can be used to obtain a good initial tessellation of the domain Ω. Using bound (26), we can calculate a global a-priori lower bound for the value of the FOSLS functional and, hence, for the error on a given tessellation $\mathcal{T}$:

$$\sum_{\tau \in \mathcal{T}} \bar{\eta}_\tau^2 \le \mathcal{G}(\mathbf{u}^h; \mathbf{f}). \tag{28}$$

If the goal is to start the adaptive algorithm with an initial tessellation $\mathcal{T}$ that is fine enough to resolve the right-hand side (see for example [10], section 5.1), one could use a-priori lower bound (26) as an indicator of where to refine the initial tessellation. The following algorithm could be used to obtain such an initial tessellation:

1. Construct a tessellation $\mathcal{T}_0$ that resolves the geometry of Ω well.
2. Given a global lower bound $\kappa > 0$ on the FOSLS functional that is acceptable, refine $\mathcal{T}_0$ adaptively until the refined tessellation $\mathcal{T}_1$ satisfies $\sum_{\tau \in \mathcal{T}_1} \bar{\eta}_\tau^2 \le \kappa$.

Note that such an algorithm can also be useful for subsequent 'initial' tessellations involved in an FMG-like process of successively decreasing values of κ. Such a 'κ-continuation' process could be useful in the computation of an initial solution approximation with some final value of κ.

The evaluation of $\bar{\eta}_\tau$ can be done by solving a local least-squares minimization problem which involves inverting the individual element stiffness matrices. Calculating $\bar{\eta}_\tau$ for every τ is, thus, comparable in computational cost to a point relaxation sweep (where all variables at a point are relaxed simultaneously). On a structured finite element mesh, there are only a small number of different element stiffness matrices (in the case of piecewise-constant coefficients), so the calculation of $\bar{\eta}_\tau$ can be done very efficiently. Also, the information needed to evaluate $\bar{\eta}_\tau$ is entirely local to an element τ, so the a-priori error estimate can be calculated in parallel.

5. A Sharp Theoretical Estimate and Convergence of an Adaptive Algorithm. In this section, we present a theoretical result regarding error reduction in an adaptive algorithm for FOSLS. We first introduce some notation.

Suppose that $\mathbf{u}^h$ is the best approximation to the solution $\mathbf{u}$ of minimization problem (4) on the current level W^h. Let $R \subset W$ be a subregion in which further refinement is considered. Define the error $\mathbf{e} = \mathbf{u}^h - \mathbf{u}$, which we decompose into 'local' and 'harmonic' parts as follows. First, let the set of local functions (with support in R) be defined by

$$W_R = \{\mathbf{v} \in W : \mathbf{v} = 0 \text{ on } R^c \equiv \Omega - R\}. \tag{29}$$

Then, define

$$\mathbf{l} := \arg\min_{\mathbf{v} \in W_R} \mathcal{G}(\mathbf{e} + \mathbf{v}; \mathbf{0}). \tag{30}$$

Now let

$$\mathbf{h} = \mathbf{e} - \mathbf{l}, \tag{31}$$

and note that $\mathbf{h} = \mathbf{e}$ on R^c. We say that $\mathbf{h}$ is locally $\mathcal{F}$-harmonic on R since

$$\mathcal{F}(\mathbf{h};\mathbf{v}) = 0, \quad \forall \mathbf{v} \in W_R. \tag{32}$$

Note in particular that $\mathbf{e} = \mathbf{h} + \mathbf{l}$ is an $\mathcal{F}$-orthogonal decomposition of $\mathbf{e}$ in the sense that

$$\mathcal{F}(\mathbf{h};\mathbf{l}) = 0. \tag{33}$$

Note also that we can equivalently define $\mathbf{h}$ as follows:

$$\mathbf{h} := \arg \min_{\mathbf{v}=\mathbf{e} \text{ on } R^c} \mathcal{G}_R(\mathbf{v};\mathbf{0}). \tag{34}$$

The harmonic $\mathbf{h}$ is the component of the error that cannot be eliminated even with infinite refinement of R. On the other hand, the local error $\mathbf{l}$ is the component of the error that can be fully resolved by infinitely refining in R and leaving the approximation in $\overline{R^c}$ fixed. Our principle theorem shows that the FOSLS functional can be used to identify large local error provided the local region is not too 'thin'.

THEOREM 5.1. *Given region $R \subset \Omega$, approximation $\mathbf{u}^h \in W^h$, and error decomposition (31), define ε by*

$$\mathcal{G}_R(\mathbf{e};\mathbf{0}) = (1-\varepsilon)\, \mathcal{G}(\mathbf{e};\mathbf{0}). \tag{35}$$

Assume that there exists $\gamma < 1-\varepsilon$ such that

$$\mathcal{G}_R(\mathbf{h};\mathbf{0}) \le \gamma\, \mathcal{G}(\mathbf{h};\mathbf{0}). \tag{36}$$

Then

$$\mathcal{G}(\mathbf{h};\mathbf{0}) \le \delta\, \mathcal{G}(\mathbf{e};\mathbf{0}), \tag{37}$$

where $\delta = \frac{\varepsilon}{1-\gamma} < 1$.

Proof. Inequalities (35) and (36) imply

$$\begin{aligned}
(1-\gamma)\, \mathcal{G}(\mathbf{h};\mathbf{0}) &= \mathcal{G}(\mathbf{h};\mathbf{0}) - \gamma\, \mathcal{G}(\mathbf{h};\mathbf{0}) \\
&\le \mathcal{G}(\mathbf{h};\mathbf{0}) - \mathcal{G}_R(\mathbf{h};\mathbf{0}) \\
&= \mathcal{G}_{R^c}(\mathbf{e};\mathbf{0}) \\
&= \mathcal{G}(\mathbf{e};\mathbf{0}) - \mathcal{G}_R(\mathbf{e};\mathbf{0}) \\
&\le \mathcal{G}(\mathbf{e};\mathbf{0}) - (1-\varepsilon)\, \mathcal{G}(\mathbf{e};\mathbf{0}) \\
&= \varepsilon\, \mathcal{G}(\mathbf{e};\mathbf{0}).
\end{aligned}$$

Hence,

$$\mathcal{G}(\mathbf{h};\mathbf{0}) \le \frac{\varepsilon}{1-\gamma}\, \mathcal{G}(\mathbf{e};\mathbf{0}),$$

which completes the proof. □

REMARK 1. *Under the assumptions of Theorem 5.1, it is always possible to choose $\gamma \le 1-\varepsilon$, with strict inequality possible when the local error $\mathbf{l}$ is nonzero.*

Proof. We have

$$\mathcal{G}_R(\mathbf{h};\mathbf{0}) \le \mathcal{G}_R(\mathbf{e};\mathbf{0}) = \frac{1-\varepsilon}{\varepsilon} \mathcal{G}_{R^c}(\mathbf{e};\mathbf{0}). \tag{38}$$

Adding to this the tautology

$$\frac{1-\varepsilon}{\varepsilon}\mathcal{G}_R(\mathbf{h};\mathbf{0}) = \frac{1-\varepsilon}{\varepsilon}\mathcal{G}_R(\mathbf{h};\mathbf{0}), \tag{39}$$

we get

$$(1+\frac{1-\varepsilon}{\varepsilon})\mathcal{G}_R(\mathbf{h};\mathbf{0}) \le \frac{1-\varepsilon}{\varepsilon}(\mathcal{G}_{R^c}(\mathbf{e};\mathbf{0}) + \mathcal{G}_R(\mathbf{h};\mathbf{0})), \tag{40}$$

or

$$\mathcal{G}_R(\mathbf{h};\mathbf{0}) \le (1-\varepsilon)\mathcal{G}(\mathbf{h};\mathbf{0}). \tag{41}$$

Hence, we can always choose γ in (36) to be less or equal to $1-\varepsilon$. Note that the inequality in (38) is strict when $\mathbf{l} \neq \mathbf{0}$, in which case we may choose $\gamma < 1-\varepsilon$. □

We have not used any specific information about $\mathcal{G}$, so Theorem 5.1 holds for any functional. We have also not yet used any information about how $\mathbf{u}^h$ was obtained, so this estimate really holds for any approximation in W.

Equality (35) can be interpreted in the following way. We want to choose a refinement region $R \subset \Omega$ that contributes a significant portion to the total value of the functional. The idea is to identify a subregion R of Ω so that ε defined by (35) is fairly small. Inequality (36) is a statement about the shape of the refinement region R. If R is too 'thin', then the constant γ will be close to 1, in which case a lot of energy (i.e., value of $\mathcal{G}$) can be hidden (inaccessible to refinement) in the $\mathcal{F}$-harmonic part of the error. Hence, we want to choose a refinement region that is not too small (nor too large or else global refinement would be more efficient). Inequality (37) means that exact solution of the minimization problem in R would reduce the total value of the functional by at least a factor of δ. The theorem thus asserts that, under certain restrictions on the refined region, the FOSLS functional can be a sharp predictor of the error that refinement would eliminate.

In practice, infinite refinement is generally impossible. To illustrate how Theorem 5.1 can be applied when only one refinement level is used in R, assume that refinement by halving h reduces the local error $\mathbf{l}$ in R by a factor of $\frac{1}{4}$:

$$\mathcal{G}(\mathbf{u}^{h/2};\mathbf{f}) \le \mathcal{G}_{R^c}(\mathbf{e};\mathbf{0}) + \mathcal{G}_R(\mathbf{h};\mathbf{0}) + \frac{1}{4}\mathcal{G}_R(\mathbf{l};\mathbf{0}). \tag{42}$$

Here $\mathbf{u}^{h/2}$ is equal to the initial approximation, $\mathbf{u}^h$, on $\bar{R}^c$, but it is obtained by solving the local FOSLS minimization problem on R. Consider the relations

$$\mathcal{G}_R(\mathbf{l};\mathbf{0}) = \mathcal{G}_R(\mathbf{e};\mathbf{0}) - \mathcal{G}_R(\mathbf{h};\mathbf{0}), \tag{43}$$

which follows from (33), and

$$\mathcal{G}_R(\mathbf{h};\mathbf{0}) \le \frac{\gamma}{1-\gamma}\mathcal{G}_{R^c}(\mathbf{e};\mathbf{0}), \tag{44}$$

which follows from (36) and the fact that $\mathbf{h} = \mathbf{e}$ on R^c and $\mathcal{G}(\mathbf{h};\mathbf{0}) = \mathcal{G}_R(\mathbf{h};\mathbf{0}) + \mathcal{G}_{R^c}(\mathbf{h};\mathbf{e})$. Then (42), (43), (44), and (35) imply

$$\begin{aligned}
\mathcal{G}(\mathbf{u}^{h/2};\mathbf{f}) &\le \mathcal{G}_{R^c}(\mathbf{e};\mathbf{0}) + \mathcal{G}_R(\mathbf{h};\mathbf{0}) + \frac{1}{4}\left(\mathcal{G}_R(\mathbf{e};\mathbf{0}) - \mathcal{G}_R(\mathbf{h};\mathbf{0})\right) \\
&= \mathcal{G}_{R^c}(\mathbf{e};\mathbf{0}) + \frac{3}{4}\mathcal{G}_R(\mathbf{h};\mathbf{0}) + \frac{1}{4}\mathcal{G}_R(\mathbf{e};\mathbf{0})
\end{aligned}$$

$$
\begin{aligned}
&\leq \left(1+\frac{3}{4}\frac{\gamma}{1-\gamma}\right)\mathcal{G}_{R^c}(\mathbf{e};\mathbf{0})+\frac{1}{4}\mathcal{G}_R(\mathbf{e};\mathbf{0})\\
&= \left(1+\frac{3}{4}\frac{\gamma}{1-\gamma}\right)\varepsilon\,\mathcal{G}(\mathbf{e};\mathbf{0})+\frac{1}{4}(1-\varepsilon)\mathcal{G}(\mathbf{e};\mathbf{0})\\
&= \left(\frac{1}{4}+\frac{3}{4}\frac{\varepsilon}{1-\gamma}\right)\mathcal{G}(\mathbf{e};\mathbf{0}).
\end{aligned}
$$

Thus, we obtain a bound similar to (37), but now with $\delta = \frac{1}{4} + \frac{3}{4}\frac{\varepsilon}{1-\gamma}$. This implies convergence for the case of one additional level of refinement when $\frac{\varepsilon}{1-\gamma} < 1$ (i.e., $\mathbf{l} \neq \mathbf{0}$).

6. Conclusions. We want to stress that the bounds on the a-posteriori error estimate η_τ follow immediately from standard FOSLS theory. This is a substantial advantage over what has been obtained for other methods (see, for example, [1], [2], or, for a review of a-posteriori error estimators, [13]). We have also showed how FOSLS naturally yields a local a-posteriori error estimate. The numerical results presented in Section 3.2 support the practicality of the proposed a-posteriori error estimate for adaptively refining a finite element mesh.

The local a-priori lower error bound $\bar{\eta}_\tau$ can be an important tool in the context of adaptive refinement, either for generating an initial mesh or for use in the process of adaptively refining a computational mesh. An important strength of this measure is that an unacceptably large value $\bar{\eta}_\tau$ is a certain signal that τ must be refined.

Theorem 5.1 shows that the FOSLS functional provides a sharp measure of the local error, provided estimate (36) holds. This restriction is related to the geometry of R and the nature of the specific FOSLS functional. To confirm (36), we need to be sure that local $\mathcal{F}$-harmonics do not exhibit inordinately large local energies (otherwise, the refinement process could be misled by large local $\mathcal{F}$-harmonic errors, which cannot be eliminated within the local region itself). What is needed is an articulation of the assumptions on $\mathcal{G}$ and specific restrictions on R that could be used to guide the refinement process. This should lead to a proof of convergence of the refinement strategy that is potentially sharper and more general than other theories (see [10] for important recent work on Poisson's equation).

REFERENCES

[1] I. Babuška and W. C. Rheinbold, *Error estimates for adaptive finite element computations*, SIAM J. Numer. Anal., 15 (1978), pp. 736–754.

[2] R. E. Bank and R. K. Smith, *A posteriori error estimates based on hierarchical bases*, SIAM J. Numer. Anal., 30 (1993), pp. 921–935.

[3] P. B. Bochev and M. D. Gunzburger, *Analysis of least squares finite element methods for the Stokes equations*, Math. Comp., 63 (1994), pp. 479–506.

[4] ———, *Least-squares methods for the velocity-pressure-stress formulation of the Stokes equations*, Comp. Meth. Appl. Mech. Eng., 126 (1995), pp. 267–287.

[5] Z. Cai, R. D. Lazarov, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for second-order partial differential equations: Part I*, SIAM J. Numer. Anal., 31 (1994), pp. 1785–1799.

[6] Z. Cai, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for velocity-vorticity-pressure form of the Stokes equations, with applications to linear elasticity*, E.T.N.A., 3 (1995), pp. 150–159.

[7] ———, *First-order system least squares for second-order partial differential equations: Part II*, SIAM J. Numer. Anal., 34 (1997), pp. 425–454.

[8] ———, *First-order system least squares for the Stokes equations with applications to linear elasticity*, SIAM J. Numer. Anal., (to appear).

[9] C. L. Chang, *An error estimate of the least squares finite element method for the Stokes problem in three dimensions*, Math. Comp., 63 (1994), pp. 41–50.

[10] W. Dörfler, *A convergent adaptive algorithm for poisson's equation*, SIAM J. Numer. Anal., 33 (1996), pp. 1106–1124.

[11] J. M. Fiard, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares (FOSLS) for convection-diffusion problems: Numerical results*, SIAM J. Sci. Comput., (to appear).

[12] B. N. Jiang and G. F. Carey, *Adaptive refinement for least-squares finite elements with element-by-element conjugate gradient solution*, Int. J. Numer. Meth. Engr., 24 (1987), pp. 569–580.

[13] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, 1996.