SIAM J. SCI. COMPUT.
Vol. 32, No. 3, pp. 1506–1526

NESTED ITERATION AND FIRST-ORDER SYSTEM LEAST SQUARES FOR INCOMPRESSIBLE, RESISTIVE MAGNETOHYDRODYNAMICS[*]

J. H. ADLER[†], T. A. MANTEUFFEL[‡], S. F. MCCORMICK[‡], J. W. RUGE[‡], AND G. D. SANDERS[‡]

Abstract. This paper develops a nested iteration algorithm to solve time-dependent nonlinear systems of partial differential equations. For each time step, Newton's method is used to form approximate solutions from a sequence of nested spaces, where the resolution of the approximations increases as the algorithm progresses. Nested iteration results in most of the iterations being performed on coarser grids, where minimal work is needed to reduce error to the level of discretization error. The approximate solution on a given coarse grid is interpolated to a refined grid and is used as an initial guess for the problem posed there. The approximation is then already close enough to the solution on the current grid that a minimal amount of work is needed to solve the refined problem due to the rapid convergence of Newton's method near a solution. The paper develops an algorithm that attempts to optimize accuracy-per-computational-cost on each grid, so that essentially no unnecessary work is done on any grid. The nested iteration algorithm is then applied to a reduced two-dimensional model of the incompressible, resistive magnetohydrodynamic (MHD) equations. Using this algorithm on the MHD equations in the context of a first-order system least squares finite element discretization and algebraic multigrid to solve the linearized systems, instabilities in a model tokamak fusion reactor are simulated. Numerical results show that this highly complex nonlinear problem is solved in an equivalent of 30–80 fine-grid relaxation sweeps per time step.

Key words. magnetohydrodynamics, algebraic multigrid, first-order system least squares, nested iteration

AMS subject classifications. 65F10, 65M55, 76W05

DOI. 10.1137/090766905

1. Introduction.

This paper applies a nested iteration algorithm to a set of incompressible resistive magnetohydrodynamic (MHD) equations. A first-order systems least squares (FOSLS) [11, 12] finite element discretization is used. The main focus of this paper is to show that nested iteration is a useful tool for solving complicated systems of nonlinear partial differential equations. Many other implicit nonlinear solvers have been developed for the MHD equations, specifically in [13, 14, 19, 20, 21, 26]. However, our aim is to describe the nested iteration strategy applied in the Newton-FOSLS framework. This is a continuation of work found in [30, 33]. Here we develop detailed stopping criteria for the steps in the nested iteration algorithm to make the algorithm yield the most accuracy-per-computational-cost. As is shown for the MHD equations, nested iteration allows for a complicated system of nonlinear equations to be solved to within discretization accuracy in only a handful of work units (WUs). A WU is defined as the amount of computation required to perform one fine-grid relaxation sweep. The main idea is that most of the work is done on coarse grids, where computation is much cheaper. The approximation to the solution on these coarse

[*] Received by the editors August 1, 2009; accepted for publication (in revised form) April 21, 2010; published electronically May 26, 2010.

http://www.siam.org/journals/sisc/32-3/76690.html

[†] Department of Applied Mathematics, University of Colorado at Boulder, Boulder, CO 80309-0526. Current address: Mathematics Department, Pennsylvania State University, University Park, PA 16802 (adler@math.psu.edu).

[‡] Department of Applied Mathematics, University of Colorado at Boulder, Boulder, CO 80309-0526 (tmanteuf@colorado.edu, stevem@colorado.edu, jruge@colorado.edu, sandersg@colorado.edu).

grids is interpolated up to successively finer grids, where linearizations and algebraic solves are applied. The process is then continued to finer grids until a desired error tolerance or resolution is reached. The result is that coarse grids give fairly accurate initial guesses for solving on the fine grids, and thus less work is needed to approximate the solution at the desired resolution. The algorithm is designed to maximize the accuracy-per-computational-cost of solving the system of equations. To accomplish this efficiently, estimates of the accuracy on the next finer grid and the work that is required to obtain that solution are needed. Thus, the discretization methods used need to give good a posteriori error measures. In addition, an estimate of how well the algebraic solver performs is needed. This includes estimates of the convergence factors of the algebraic solver and how much more work it costs to perform such an iteration on a fine grid compared to a coarse grid. The algorithm described in this framework can be applied to any discretization and linear solver method that produces these estimates. Thus, in this sense, it is designed in a general framework. However, the specific application described is MHD, FOSLS is the discretization formulation that is used, and algebraic multigrid (AMG) is the linear solver that is applied. The a posteriori error measures make it very amenable to the nested iteration algorithm that comes naturally with FOSLS. As a result, the resistive MHD system described in section 4 is solved to within discretization error using only 30–80 WUs per time step.

The MHD system and the FOSLS formulation applied to it are discussed in detail in the companion paper [3], so we include only a brief description here in section 2. In section 3, the nested iteration algorithm is developed. An approach that attempts to minimize the number of linearization steps and AMG V-cycles required to obtain an approximation within discretization error on the current grid is described. A strategy has been developed for this, parts of which are described in [2, 33]. We go into more detail here. In section 4, numerical results are shown for a two-dimensional (2D) reduced model that simulates plasma instabilities in a tokamak reactor. These results confirm that the nested iteration approach yields solutions to the test problems in only a handful of WUs.

2. The MHD equations and FOSLS formulation. The resistive MHD equations are time-dependent and nonlinear and involve several dependent variables. The system is a coupling of the incompressible Navier–Stokes and Maxwell's systems. The primitive variables are defined to be the fluid velocity, $\mathbf{u}$, the fluid pressure, p, the magnetic field, $\mathbf{B}$, the current density, $\mathbf{j}$, and the electric field, $\mathbf{E}$. In addition, a resistive form of Ohm's law,

$$\mathbf{j} = \sigma(\mathbf{E} + \mathbf{u} \times \mathbf{B}), \tag{2.1}$$

is used to eliminate the electric field, $\mathbf{E}$, from the equations. After a nondimensionalization using Alfvén units, the following equations for incompressible resistive MHD are obtained (i.e., Navier–Stokes coupled with Maxwell's equations) [23, 32]:

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} - \mathbf{j} \times \mathbf{B} + \nabla p - \frac{1}{R_e}\nabla^{\mathbf{2}} \mathbf{u} = \mathbf{f}, \tag{2.2}$$

$$\frac{\partial \mathbf{B}}{\partial t} - \mathbf{B} \cdot \nabla \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{B} + \frac{1}{S_L}(\nabla \times \mathbf{j}) = \mathbf{g}, \tag{2.3}$$

$$\nabla \times \mathbf{B} = \mathbf{j}, \tag{2.4}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{2.5}$$

$$\nabla \cdot \mathbf{u} = 0, \tag{2.6}$$

$$\nabla \cdot \mathbf{j} = 0. \tag{2.7}$$

Here R_e is the fluid Reynolds number, and S_L is the Lundquist number, both of which are assumed to be constants and adjusted for different types of physical behavior.

Using the FOSLS method [11, 12], the system is first put into a differential first-order system of equations. This is done based on a vorticity-velocity-pressure-current formulation. Since explicit vorticity boundary conditions are provided in all the test problems, this is an appropriate formulation. A scaling analysis is performed in [3], which yields a nice block structure for the MHD system. This results in good AMG convergence of the linear systems obtained, while still preserving the physics of the system.

Vorticity, $\boldsymbol{\omega} = \nabla \times \mathbf{u}$, is introduced, and the final formulation in three dimensions is

$$\frac{1}{\sqrt{R_e}}\nabla \times \mathbf{u} - \sqrt{R_e}\boldsymbol{\omega} = 0, \tag{2.8}$$

$$\frac{1}{\sqrt{R_e}}\nabla \cdot \mathbf{u} = 0, \tag{2.9}$$

$$\sqrt{R_e}\nabla \cdot \boldsymbol{\omega} = 0, \tag{2.10}$$

$$\frac{1}{\sqrt{R_e}}\frac{\partial \mathbf{u}}{\partial t} - \mathbf{u} \times \boldsymbol{\omega} - \mathbf{j} \times \mathbf{B} - \sqrt{R_e}\nabla p + \frac{1}{\sqrt{R_e}}\nabla \times \boldsymbol{\omega} = \mathbf{f}, \tag{2.11}$$

$$\frac{1}{\sqrt{S_L}}\nabla \times \mathbf{B} - \sqrt{S_L}\mathbf{j} = 0, \tag{2.12}$$

$$\frac{1}{\sqrt{S_L}}\nabla \cdot \mathbf{B} = 0, \tag{2.13}$$

$$\sqrt{S_L}\nabla \cdot \mathbf{j} = 0, \tag{2.14}$$

$$\frac{1}{\sqrt{S_L}}\frac{\partial \mathbf{B}}{\partial t} + \frac{1}{\sqrt{R_e S_L}}(\mathbf{u} \cdot \nabla \mathbf{B} - \mathbf{B} \cdot \nabla \mathbf{u}) + \frac{1}{\sqrt{S_L}}\nabla \times \mathbf{j} = \mathbf{g}. \tag{2.15}$$

An implicit backward differencing formula (BDF) time-stepping method is applied to this system to yield a nonlinear system of equations, $\mathcal{L}(u) = f$, where $\mathcal{L}$ is closely related to the left-hand side of (2.8)–(2.15) and f is a vector of known functions. Here $u = (\mathbf{u}, \boldsymbol{\omega}, p, \mathbf{B}, \mathbf{j})^T$ represents a vector of all of the dependent variables that should not be confused with the vector fluid velocity, $\mathbf{u}$. Both [2] and [3] discuss the stability and convergence of the time–stepping on the MHD problem using FOSLS. Next the L^2 norm of the residual of this system is minimized. This is a nonquadratic functional, but since it is the functional of the nonlinear operator, we refer to it, for simplicity, as the *nonlinear functional*:

$$\mathcal{F}(u) = ||\mathcal{L}(u) - f||_0. \tag{2.16}$$

In general, we wish to find the arg min of (2.16) in some solution space $\mathcal{V}$. We choose $\mathcal{V}$ to be an H^1 product space with boundary conditions that are chosen to satisfy the physical constraints of the problem as well as the assumptions needed for the FOSLS framework. In practice, a series of nested finite subspaces, $\mathcal{V}^h$, are used to approximate the solution in $\mathcal{V}$. However, in the Newton-FOSLS approach [16, 15], system (2.8)–(2.15) is first linearized using a Newton step, thus forming a linear problem and an associated functional to minimize. This so-called Gauss–Newton method [18] is an alternative to applying Newton's method directly to the nonlinear variational system of (2.16). The linearization is done about some approximation, u_i^h, in the finite element space, $\mathcal{V}^h$, and the change, δ_i^h, is found such that the next iterate is $u^h_{i+1} = u^h_i + \delta^h_i$. The linearized equation is then written as

$$\mathcal{L}'(u^h_i)[\delta^h_i] = f - \mathcal{L}(u^h_i). \tag{2.17}$$

Here, $\mathcal{L}'(u^h_i)[\delta^h_i]$ is the Fréchet derivative of $\mathcal{L}$ in the direction δ^h_i. The well-posedness of this linearized problem is shown by proving H^1 ellipticity of the linearized system. This has been shown in [2] under certain assumptions, but we omit such results here in the interest of brevity.

The linearized operator, $\mathcal{L}'$, is a mapping from $\mathcal{V}^h$ to an L^2 product space. This linear problem is now recast as the minimization of a functional constructed by taking the L^2 norm of the residual of each equation in the linearized system. This is written as

$$\delta^h_* = \arg\min_{\delta^h \in \mathcal{V}^h} F(\delta^h) := \arg\min_{\delta^h \in \mathcal{V}^h} ||\mathcal{L}'(u^h_i)[\delta^h] - (f - \mathcal{L}(u^h_i))||_0, \tag{2.18}$$

where δ^h_* is the solution in $\mathcal{V}^h$. Here $F(\delta^h)$ is actually a quadraticized functional, but since it is the functional of the linearized operator, it is from here onword referred to as the linearized functional. The minimization of this functional results in the weak form of the problem that produces a symmetric positive definite algebraic system for each Newton step. The solution, δ^h_*, satisfies the bilinear weak form:

$$\left\langle \mathcal{L}'(u^h_i)[\delta^h_*], \mathcal{L}'(u^h_i)[v^h] \right\rangle = \left\langle f - \mathcal{L}(u^h_i), \mathcal{L}'(u^h_i)[v^h] \right\rangle \qquad \forall v^h \in \mathcal{V}^h. \tag{2.19}$$

Any element in the finite element space, v^h, can be decomposed into the basis function elements, ϕ_j. Thus,

$$v^h = \sum_j \alpha_j \phi_j.$$

Then the components of the algebraic linear system, $Ax = b$, are as follows:

$$A_{ij} = \left\langle \mathcal{L}'(u^h_i)[\phi_j], \mathcal{L}'(u^h_i)[\phi_i] \right\rangle, \tag{2.20}$$

$$b_i = \left\langle f - \mathcal{L}(u^h_i), \mathcal{L}'(u^h_i)[\phi_i] \right\rangle. \tag{2.21}$$

In addition, proving continuity and coercivity of the resulting bilinear form results in H^1 equivalence of the FOSLS functional. In other words, there exists positive constants c_0 and c_1 such that

$$c_0||\delta^h||_1 \leq F(\delta^h) \leq c_1||\delta^h||_1 \qquad \forall \delta^h \in \mathcal{V}^h. \tag{2.22}$$

Moreover, the FOSLS functional yields a sharp a posteriori local error estimate, which is used to make the algorithm more robust. Additionally, under the right conditions, the FOSLS framework produces algebraic systems that are solved easily by multilevel iterative solvers. Our choice here is the AMG [7, 6, 5, 10, 25, 28], which, when applied to the FOSLS discretization, has been shown to be an optimal $(O(n))$ solver for many example problems [10, 11, 12]. As stated above, using the formulation in (2.8)–(2.15), and with appropriate boundary conditions, H^1 equivalence of the linearized FOSLS functional is shown in [2]. Therefore, the FOSLS functionals are a good measure of the H^1 norm of the global error and yield a seminorm of the local error. Thus, they can be used to develop an efficient solution algorithm.

3. Nested iteration strategies.

The goal for this section is to establish criteria to control the number of Newton steps and linear iterations employed at each refinement level. The idea is to solve the problem with the most accuracy-per-computational-cost. An inexact Gauss–Newton method embedded in a nested iteration approach was presented before in the context of nonlinear first-order system least squares problems in [30]. What is new here is the detailed derivation of termination criteria for the various subproblems on different refinement levels in order to optimize the performance with respect to overall work. Heuristic ideas of balancing the overall work in a multilevel framework for inexact Newton multilevel finite element methods was also described in [18]. The results below show that only a few Newton steps and V-cycles are needed to get a good approximation to the solution on the finest grid. Under natural hypotheses, which are in force here, this was proved in [16, 15]. In addition, we discuss the heuristic that it is not necessary to solve the discrete problems on these coarse grids much below the level of discretization error. In fact, at some point, doing more work on the coarser grid is not as efficient as just moving to a finer grid. It is at this point that the FOSLS methodology proves to be important. Using the a posteriori error estimate of the functional, we determine how well both the linear and nonlinear systems are being solved. We refer to F and $\mathcal{F}$ as the linearized and nonlinear functionals, respectively, as in (2.18) and (2.16).

With the linearized functional, along with the AMG convergence factors, an estimate of how much the error of the linear system is being reduced and at what cost is computed. Likewise, by measuring the nonlinear functional and by determining the number of Newton steps performed, the accuracy-per-computational-cost required to solve the nonlinear system is estimated. These estimates are then used to establish stopping parameters for the linear iterations as well as for the Newton iterations. For the linear systems, we decide if the best strategy is to either perform another V-cycle on the current grid or to relinearize the system. Work here is an extension of [2, 33], where the iterations were stopped when the functional value of the current iterate came within a given tolerance of the functional minimum. The Newton iterations are also controlled in this manner, by determining how close the approximate solution is to the discrete solution of the nonlinear problem on a given grid. If certain criteria are met, no more linearizations are performed and the computation is transferred to a finer grid. In the next two subsections, we describe the stopping criteria in more detail.

3.1. The linear problem.

First, assume that the system has already been linearized around some point u_0 and that only a linear problem, $L\delta = g$, is being solved. In the full nonlinear setting, $L = \mathcal{L}'$, δ is the change from the Newton step and $g = f - \mathcal{L}(u_0)$. Assume that there are only two grids, grid h and grid $2h$, and, for simplicity, that uniform refinement is used to transfer between them. In practice, due to the local error measures that FOSLS provides, an adaptive mesh refinement (AMR) method can be employed as well. This is done in the proceedings paper, [1]. Using uniform refinement, though, moving to a finer grid increases the number of elements on the finite element mesh by 2^d, where d is the dimension of the problem. Let the true continuous solution to the problem be δ_*, and let δ_*^{2h} and δ_*^h be the discrete solutions on their respective grids. Since the FOSLS method is used, the functional is an estimate of the error. We denote the discretization error, measured in the functional norm, on grids h and $2h$ by F_*^h and F_*^{2h}, respectively. In the process of solving on grid $2h$, several iterates, δ_i^{2h}, $i = 0, 1, \ldots$, are obtained. At iteration i, with iterate δ_i^{2h}, the error is measured and denoted by F_i^{2h}. Since the error is orthogonal to the finite element space in the functional norm, as can be seen from (2.19), it is convenient to write this as a multiple of the discretization error on grid $2h$; that is,

$$(F_i^{2h})^2 = (1 + (\epsilon_i^{2h})^2)(F_*^{2h})^2. \tag{3.1}$$

This is also seen in Figure 3.1. The value $\epsilon_i^{2h} F_*^{2h}$ is referred to as the algebraic error on grid $2h$ with ϵ_i^{2h} being the fraction of the discretization error. Also, assume that the algebraic solver has a steady convergence factor that is independent of the size of the grid or the current iterate. This is not precisely true, but the convergence factor is assumed to be bounded above by a constant, $\rho < 1$, independent of the mesh. Thus, we assume that ρ is known and that $\epsilon_{i+1}^{2h} = \rho \epsilon_i^{2h}$. In practice, an estimate of ρ is obtained from a sequence of residuals of the matrix equations.

Figure 3.1 is intended to demonstrate the strategy for determining when to stop iterating on grid $2h$ and move to grid h. Suppose that iterations are done on grid $2h$ and that the current iterate is δ_s^{2h}, with algebraic error $\epsilon_s^{2h} F_*^{2h}$. Grid $2h$ is indicated in Figure 3.1 by the subspace restricted to the x-axis, while the plane represents grid h and the true solution, δ_*, is shown hovering above δ_*^h. The discrete solution on grid h is closer to the true solution than the coarse grid approximation. In addition, iterating on grid $2h$ restricts us to only getting closer to δ_*^{2h}. Therefore, better accuracy is obtained by iterating on the finer grid, h. However, this is more costly, so it is important to know how much accuracy is gained by doing one more iteration on grid $2h$ and by comparing this to the amount of accuracy gained by interpolating to grid h and moving directly toward δ_*^h. That is, from the position denoted by δ_s^{2h}, is it better to move to position δ_{s+1}^{2h} or to the position denoted by δ_1^h? Of course, a measure of the accuracy-per-computational-cost must be established for both of these choices. By assumption, computation on grid h costs 2^d times the cost of an iteration on grid $2h$, and, by examining the triangles in Figure 3.1, the reduction of the two errors from the two types of iteration is determined. In the next few paragraphs, we quantify a measure that helps us determine an optimal strategy.

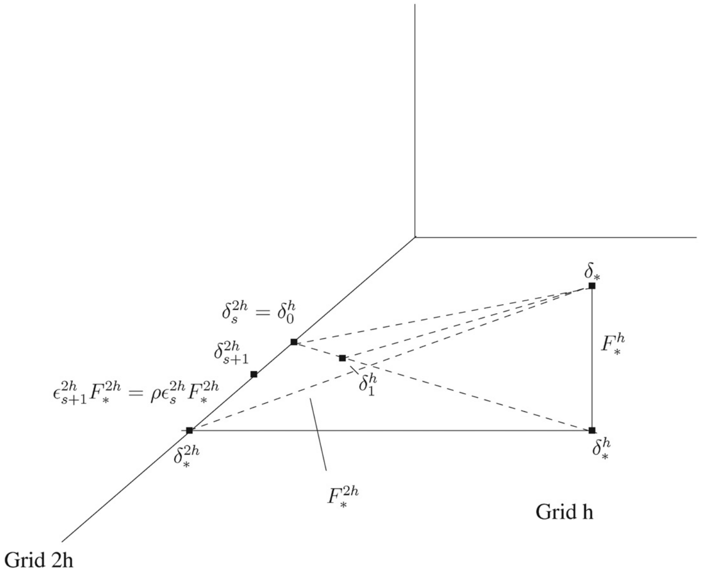

FIG. 3.1. *Linear iterations along grid $2h$ and grid h.*

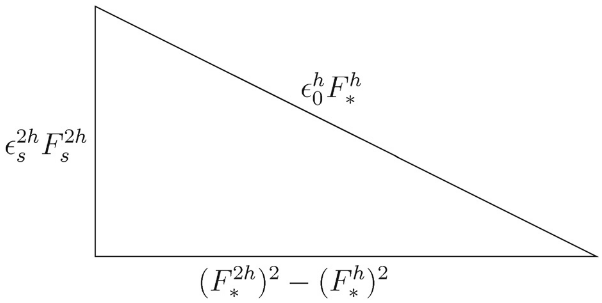

FIG. 3.2. *Relationship of algebraic errors on different grids.*

Staying on grid $2h$ and iterating once more yields the new total error for iterate δ_{s+1}^{2h} as a sum of the discretization error plus the algebraic error, which is approximated as

$$F_{s+1}^{2h} = ((F_*^{2h})^2 + (\rho\epsilon_s^{2h} F_*^{2h})^2)^{\frac{1}{2}}. \tag{3.2}$$

If, instead, the approximate solution is moved to grid h and an iteration is performed to obtain u_1^h, then the error can be approximated as $F_1^h = ((F_*^h)^2 + (\rho\epsilon_0^h F_*^h)^2)^{\frac{1}{2}}$. However, to compare these two measures, we must account for the fact that an iteration on grid h is more expensive. Instead of looking at the regular convergence factor, ρ, we look at the effective convergence factor, $\rho^{1/\mathrm{Work}}$; thus, the effective error on grid h is approximated as

$$F_1^h = ((F_*^h)^2 + (\rho^{\frac{1}{2^d}}\epsilon_0^h F_*^h)^2)^{\frac{1}{2}}. \tag{3.3}$$

The two measures are equivalent when

$$(F_*^{2h})^2 + (\rho\epsilon_s^{2h} F_*^{2h})^2 = (F_*^h)^2 + \rho^{2^{1-d}}(\epsilon_0^h F_*^h)^2. \tag{3.4}$$

Extracting the triangle in Figure 3.2 from Figure 3.1, the algebraic error on grid $2h$, $\epsilon_s^{2h} F_*^{2h}$, can be related to the algebraic error on grid h, $\epsilon_0^h F_*^h$. Then, after some algebra, we obtain

$$(\hat{\epsilon}^{2h})^2 = \frac{\left(\rho^{2^{1-d}} - 1\right)\left(1 - \left(\frac{F_*^h}{F_*^{2h}}\right)^2\right)}{\rho^2 - \rho^{2^{1-d}}}. \tag{3.5}$$

In other words, in terms of accuracy-per-computational-cost, it is more efficient to move to the finer grid when the algebraic error on grid $2h$ is less than $\hat{\epsilon}^{2h} F_*^{2h}$.

Of course, this assumes that the discretization errors, F_*^{2h} and F_*^h, are known. Only the total error, or linearized FOSLS functional value, of the current iterate on grid $2h$ is known, though. One way to estimate the discretization errors is to assume that the algebraic error decreases at a linear rate. Since the functional squared is a quadratic, an estimate of the discretization error, or the minimum of the quadratic functional, can be extrapolated. We assume

$$(F_s^{2h})^2 = (1 + (\epsilon_s^{2h})^2)(F_*^{2h})^2$$

and that

$$
\begin{aligned}
(F_{s+1}^{2h})^2 &= (1+(\epsilon_{s+1}^{2h})^2)(F_*^{2h})^2 \\
&= (1+\rho^2(\epsilon_s^{2h})^2)(F_*^{2h})^2.
\end{aligned}
$$

Thus, three consecutive steps yield an estimate of both ϵ_s^{2h} and F_*^{2h}. This may require doing a few extra iterations to get a good estimate of the minimum, but since this work is done on coarse grids, we argue that this does not increase the total work significantly.

Another approach to stopping the linear iterations involves using the discrete linear system, $Ax = b$, as in (2.20) and (2.21). Note that the main idea for solving the linear system is that the iterations should stop once the algebraic error is less than $\hat{\epsilon}^{2h}$ of the discretization error, where $\hat{\epsilon}^{2h}$ is as in (3.5). Therefore, the stopping criterion we seek is

$$
||L\delta_i^{2h} - L\delta_*^{2h}||_0 \le \hat{\epsilon}^{2h} F_*^{2h}. \tag{3.6}
$$

Using (3.1), this can be rewritten as

$$
||L\delta_i^{2h} - L\delta_*^{2h}||_0 \le \frac{\hat{\epsilon}^{2h}}{\sqrt{1+(\hat{\epsilon}^{2h})^2}} F_i^{2h} := \tilde{\epsilon}^{2h} F_i^{2h}. \tag{3.7}
$$

The left-hand side is in the range of L applied to $\mathcal{V}^{2h}$, so (3.7) is equivalent to

$$
\sup_{v^{2h}\in\mathcal{V}^{2h}} \frac{\left\langle L\delta_i^{2h} - L\delta_*^{2h}, Lv^{2h}\right\rangle}{||Lv^{2h}||_0} \le \tilde{\epsilon}^{2h} F_i. \tag{3.8}
$$

Since δ_*^{2h} is the minimum of the linearized functional, the Fréchet derivative of the square of the functional, $G(\delta_*^{2h})$, is zero in all directions:

$$
G'(\delta_*^{2h})[v^{2h}] = 2\left\langle L\delta_*^{2h} - g, Lv^h\right\rangle = 0 \qquad \forall v^h \in \mathcal{V}^h. \tag{3.9}
$$

Thus, $\left\langle L\delta_*^{2h}, Lv^{2h}\right\rangle = \left\langle g, Lv^{2h}\right\rangle$, and (3.8) becomes

$$
\sup_{v^{2h}\in\mathcal{V}^{2h}} \frac{\left\langle L\delta_i^{2h} - g, Lv^{2h}\right\rangle}{||Lv^{2h}||_0} \le \tilde{\epsilon}^{2h} F_i. \tag{3.10}
$$

In addition, using the fact that the linearized operator is H^1 elliptic, we use the coercivity bound, as in (2.22), and the fact that $||v^{2h}||_0 \le ||v^{2h}||_1$ to get that (3.10) is satisfied if the following is true:

$$
\sup_{v^{2h}\in\mathcal{V}^{2h}} \frac{\left\langle L\delta_i^{2h} - g, Lv^{2h}\right\rangle}{||v^{2h}||_0} \le \tilde{\epsilon}^{2h} c_0 F_i. \tag{3.11}
$$

The bound (3.11) can be replaced by a bound in terms of quantities associated with the solution of the algebraic equations. Using (2.20) and (2.21) and defining the algebraic residual as $r_i = b - Ax_i$, then (3.11) is satisfied when

$$
(r_i^T r_i)^{\frac{1}{2}} \le \frac{\tilde{\epsilon}^{2h} c_0}{||M^{-1}||} (x_i^T A x_i - 2b^T x_i + ||g||_0^2)^{1/2}. \tag{3.12}
$$

Here x_i is the current iterate of the coefficient vector for the solution to the linear system, and $||M^{-1}||$ is the l^2 norm of the inverse of the mass matrix associated with $\mathcal{V}^h$.

If this stopping criterion is met, then (3.7) is satisfied. In many applications, these quantities are readily available. However, if they are difficult to obtain, then estimating F_*^{2h} as described above is a viable alternative and is used in our numerical results.

In either case, whether using (3.1) or (3.12) to obtain (3.7), the convergence factor, ρ, and the reduction factor due to the refinement, F_*^h/F_*^{2h}, are needed. Information from a previous grid can be used. For instance, when uniform refinement is used, F_*^h is approximated from F_*^{2h} using the assumption that

$$F_*^h = \frac{1}{2^p} F_*^{2h},$$

where p is the order of the finite element space used. In [1, 17, 24], an efficiency-based AMR scheme, known as ACE, is used that estimates the actual error reduction in each element and determines whether more refinement is necessary. This can then be used to determine the reduction factor in this setting. Thus, we have shown how to determine when to perform another linear iteration on a given grid in the case of the linearized functional. In the next section, we discuss the overall nonlinear problem and derive stopping criteria that indicate when the nonlinear functional is solved on the order of discretization error associated with each grid.

3.2. The nonlinear problem.

The overall objective is to minimize the nonlinear functional, $\mathcal{F}^h$, on grid h. The linear solves, though, are used to minimize the linearized functional, which are assumed to be getting closer to the nonlinear functional with every Newton step. Linearization of the operator yields a quadratic minimization problem as depicted in Figure 3.3. Due to nested iteration, after a few linearizations by a quadratic functional, the minimum of the linearized functional approaches the minimum of the nonlinear functional. Therefore, comparing the linear and nonlinear functionals gives an indication of when to stop the Newton iterations. Each Newton step yields an approximation that is "close" to the predicted minimum of the linearized functional based on the stopping criteria established in subsection 3.1. If the linearized functional is "close" to the nonlinear functional, this is a good indicator that the approximation is also "close" to the minimum of the nonlinear functional. However, we need to be more precise in order to establish effective stopping criteria.

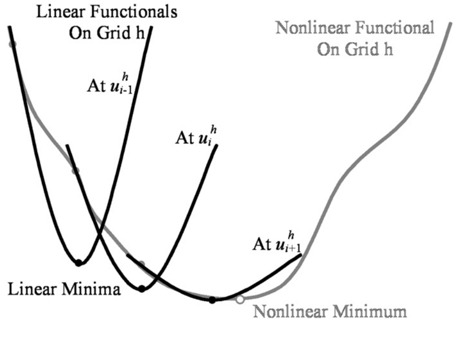

FIG. 3.3. *A one-dimensional interpretation of the nonlinear functional (in gray), compared to several iterates of the linearized functional (in black).*

The MHD system, (2.8)–(2.15), is polynomial, and the third Fréchet derivative of the nonlinear operator, $\mathcal{L}'''$, is identically zero. Let $\delta_i^h = u_{i+1}^h - u_i^h$ denote a step toward the minimum, u_*^h, from the approximation, u_i^h, and assume that the approximations are all contained in a finite element space denoted by the superscript

h. Thus, the Taylor series for this problem is finite, and the nonlinear operator is rewritten as

$$\mathcal{L}(u_i^h + \delta_i^h) = \mathcal{L}(u_i^h) + \mathcal{L}'(u_i^h)[\delta_i^h] + \frac{1}{2}\mathcal{L}''(u_i^h)[\delta_i^h, \delta_i^h]. \tag{3.13}$$

Taking the norm of the difference of the nonlinear residual from (3.13) and the linearized residual from (2.17) yields

$$F_d[\delta_i^h] := ||\frac{1}{2}\mathcal{L}''(u_i^h)[\delta_i^h, \delta_i^h]||_0. \tag{3.14}$$

Thus, the difference gives a measure of the second Fréchet derivative of the nonlinear operator in the direction of the Newton step. It is shown below that if this is small and if the linearized problem has been solved to within a fraction of discretization error, then the approximation is close to the discrete solution of the full nonlinear problem on that grid. The stopping criterion for the Newton iterations is then

$$\frac{F_d[\delta_i^h]}{\mathcal{F}(u_i^h + \delta_i^h)} \le \eta. \tag{3.15}$$

This also implies that the relative difference between the linearized and nonlinear functionals is small,

$$\frac{|F(u_{i+1}^h) - \mathcal{F}(u_{i+1}^h)|}{\mathcal{F}(u_{i+1}^h)} \le \frac{F_d[\delta_{i+1}^h]}{\mathcal{F}(u_{i+1}^h)} \le \eta \Rightarrow F(u_{i+1}^h) \le (1+\eta)\mathcal{F}(u_{i+1}^h). \tag{3.16}$$

Now, just as in the linear case, any iterations on a given grid should be stopped once the algebraic error is within a certain limit of the error. At this point, it is more efficient to interpolate to a finer grid and continue from there. Thus, just as in (3.7) for the linear problem, we want

$$||\mathcal{L}(u_{i+1}^h) - \mathcal{L}(u_*^h)||_0 \le \tilde{\epsilon}^h \mathcal{F}(u_{i+1}^h). \tag{3.17}$$

First there are several assumptions that are needed. Assume that all approximations on previous grids have been solved to within discretization error on that grid, which implies

$$\mathcal{F}(u_{i+1}^h) = O(h^p).$$

Here h is the grid parameter, and p is the order of the finite element space. By solving the linearized problem, we satisfy (3.11). It is easily concluded from this equation that

$$||\mathcal{L}'(u_i^h)[\delta_i^h]||_0 \le (1+\eta)\tilde{\epsilon}^h \mathcal{F}(u_{i+1}^h) \le (1+\eta)\tilde{\epsilon}^h O(h^p).$$

Using coercivity of the linearized functional, this also implies

$$||\delta_i^h||_0 \le ||\delta_i^h||_1 \le \frac{\tilde{\epsilon}^h}{c_0} O(h^p).$$

In addition, for the MHD problem, $\mathcal{L}''(u_i^h)[\delta_i^h, \delta_i^h] = \mathcal{L}''[\delta_i^h, \delta_i^h]$ is independent of u_i^h, and it is easy to show that for any two directions, w and z, there exists a positive constant, c_2, such that

$$||\mathcal{L}''[w,z]||_0 \le c_2 ||w||_0 ||z||_0.$$

We now use all of these estimates to get a bound on the left-hand side of (3.17).

Using the mean value theorem, the difference of the nonlinear operator applied to two approximations yields

$$\mathcal{L}(u^h_{i+1}) - \mathcal{L}(u^h_*) = \mathcal{L}'(\hat{u}^h)[u^h_{i+1} - \hat{u}^h], \tag{3.18}$$

where $\hat{u}^h = \alpha u^h_{i+1} + (1-\alpha)u^h_*$ for some $\alpha \in (0,1)$. Thus, the left-hand side of (3.17) is in the range of $\mathcal{L}'(\hat{u}^h)$ and can be written as

$$\sup_{v^h \in \mathcal{V}^h} \frac{\left\langle \mathcal{L}(u^h_{i+1}) - \mathcal{L}(u^h_*), \mathcal{L}'(\hat{u}^h)[v^h] \right\rangle}{||\mathcal{L}'(\hat{u}^h)[v^h]||_0}. \tag{3.19}$$

Since u^h_* is the discrete solution to the nonlinear problem, the Fréchet derivative of the nonlinear functional squared at this point, $\mathcal{G}(u^h_*)$, is zero in all directions:

$$\left\langle \mathcal{L}(u^h_*) - f, \mathcal{L}'(u^h_*)[v^h] \right\rangle = 0 \qquad \forall v^h \in \mathcal{V}^h. \tag{3.20}$$

By adding and subtracting some terms, (3.19) can be broken into four components:

$$\begin{aligned} \sup_{v^h \in \mathcal{V}^h} \Bigg(& \frac{\left\langle \mathcal{L}(u^h_{i+1}) - f, \mathcal{L}'(u^h_i)[v^h] \right\rangle}{||\mathcal{L}'(\hat{u}^h)[v^h]||_0} + \frac{\left\langle \mathcal{L}(u^h_{i+1}) - f, \mathcal{L}'(\hat{u}^h)[v^h] - \mathcal{L}'(u^h_i)[v^h] \right\rangle}{||\mathcal{L}'(\hat{u}^h)[v^h]||_0} \\ & + \frac{\left\langle \mathcal{L}(u^h_*) - f, \mathcal{L}'(u^h_*)[v^h] \right\rangle}{||\mathcal{L}'(\hat{u}^h)[v^h]||_0} + \frac{\left\langle \mathcal{L}(u^h_*) - f, \mathcal{L}'(\hat{u}^h)[v^h] - \mathcal{L}'(u^h_*)[v^h] \right\rangle}{||\mathcal{L}'(\hat{u}^h)[v^h]||_0} \Bigg). \end{aligned} \tag{3.21}$$

(3.21) is bounded above by taking the sup over each term. The third term here is zero from (3.20). By using the mean value theorem on $\mathcal{L}'$, as in (3.18), and the coercivity bound, the fourth term is bounded in the following way:

$$\begin{aligned} \sup_{v^h \in \mathcal{V}^h} \frac{\left\langle \mathcal{L}(u^h_*) - f, \mathcal{L}'(\hat{u}^h)[v^h] - \mathcal{L}'(u^h_*)[v^h] \right\rangle}{||\mathcal{L}'(\hat{u}^h)[v^h]||_0} & \le \frac{\mathcal{F}(u^h_*)||\mathcal{L}''[\hat{u}^h - u^h_*, v^h]||_0}{||\mathcal{L}'(\hat{u}^h)[v^h]||_0} \\ & \le \frac{\mathcal{F}(u^h_*)c_2||\hat{u}^h - u^h_*||_0}{c_0} \le \frac{c_2}{c_0} O(h^p)\mathcal{F}(u^h_*). \end{aligned}$$

The second term is bounded in the same way, replacing $\mathcal{F}(u^h_*)$ with $\mathcal{F}(u^h_{i+1})$. This leaves only the first term. Expanding this term using the full MHD operator yields

$$\sup_{v^h \in \mathcal{V}^h} \frac{\left\langle \mathcal{L}(u^h_i) - f + \mathcal{L}'(u^h_i)[\delta^h_i] + \frac{1}{2}\mathcal{L}''(u^h_i)[\delta^h_i, \delta^h_i], \mathcal{L}'(u^h_i)[v^h] \right\rangle}{||\mathcal{L}'(\hat{u}^h)[v^h]||_0}. \tag{3.22}$$

This can also be broken into two terms to get

$$\sup_{v^h \in \mathcal{V}^h} \frac{\left\langle \mathcal{L}(u^h_i) - f + \mathcal{L}'(u^h_i)[\delta^h_i], \mathcal{L}'(u^h_i)[v^h] \right\rangle}{||\mathcal{L}'(\hat{u}^h)[v^h]||_0} + \sup_{v^h \in \mathcal{V}^h} \frac{\left\langle \frac{1}{2}\mathcal{L}''[\delta^h_i, \delta^h_i], \mathcal{L}'(u^h_i)[v^h] \right\rangle}{||\mathcal{L}'(\hat{u^h})[v^h]||_0}. \tag{3.23}$$

The first term gives the exact stopping criteria for the linear iterations as in (3.10). We have assumed that this term is small. Then, using the stopping criteria, (3.15), and the estimates above, the second term can be bounded as well. Skipping most of the algebra, it can be shown that

$$\sup_{v^h \in \mathcal{V}^h} \frac{\left\langle \frac{1}{2}\mathcal{L}''[\delta_i^h, \delta_i^h], \mathcal{L}'(u_i^h)[v^h] \right\rangle}{||\mathcal{L}'(\hat{u}^h)[v^h]||_0} \leq \eta \mathcal{F}(u_{i+1}^h) \left(\frac{O(h^p)}{c_0} + 1 \right). \tag{3.24}$$

Now, with all four terms bounded, we have exactly what we want for the algebraic error of the nonlinear system. Namely, using (3.1) and (3.16), this can all be bounded in terms of the current nonlinear functional:

$$\begin{aligned} ||\mathcal{L}(u_{i+1}^h) - \mathcal{L}(u_*^h)||_0 &= \sup_{v^h \in \mathcal{V}^h} \frac{\left\langle \mathcal{L}(u_{i+1}^h) - \mathcal{L}(u_*^h), \mathcal{L}'(u_{i+1}^h)[v^h] \right\rangle}{||\mathcal{L}'(u_{i+1}^h)[v^h]||_0} \\ &\leq \left(\tilde{\epsilon}^h (1+\eta) + \eta \left(\frac{O(h^p)}{c_0} + 1 \right) + \frac{c_2}{c_0} O(h^p) + \frac{c_2 \tilde{\epsilon}^h}{\hat{\epsilon}^h c_0} O(h^p) \right) \mathcal{F}(u_{i+1}^h). \qquad (3.25) \end{aligned}$$

All of the terms involving $O(h^p)$ are assumed to be small relative to the other terms. This will eventually be valid, even if the coercivity constant, c_0, is small.

Therefore, if we are near the linearized discrete solution and the difference between the linear and nonlinear operators is small, we know that the current Newton step approximation is within a small fraction of the discrete solution to the full nonlinear problem. Then, by ensuring that the linear, (3.1), and the nonlinear, (3.15), stopping criteria are satisfied, more accuracy-per-computational-cost is gained by moving to a finer grid.

3.3. Solution algorithm.

With the pieces above, an algorithm is now devised to solve a system of nonlinear equations, $\mathcal{L}(u) = f$. Starting on a coarse grid with a given initial guess, the system is linearized and the linearized FOSLS functional is then minimized on a finite element space. At this point, several AMG V-cycles are performed until there is little to gain in accuracy-per-computational-cost. The system is then relinearized, and the minimum of the new linearized FOSLS functional is approximated in the same manner. After each set of linear solves, the relative difference between the linearized operator and the nonlinear operator is checked. If this is small, then we are close enough to the minimum of the nonlinear functional. Next the approximation is interpolated to a finer grid, where the problem is solved in the same way. This can be done adaptively or uniformly. For this paper, we show only the results using uniform refinement. However, an efficiency-based AMR method, ACE [17, 24], has been employed with success on the MHD problem. Preliminary results can be found in [1]. In any case, this process is repeated until an acceptable error has been reached or until we have run out of computational resources such as memory. If, as in the case of the MHD equations, it is a time-dependent problem, the whole process is performed at each time step. This algorithm is summarized in the flow chart in Figure 3.4.

4. Numerical results.

In this section, we investigate several MHD test problems to show that the nested iteration Newton-FOSLS-AMG method is capable of solving complex nonlinear systems in about 30–80 WUs or fine-grid relaxation equivalents. The full algorithm, as shown in Figure 3.4, was applied to two tokamak test problems [13, 14, 26, 31]. From the papers by Chacon, Knoll, and Finn [13] and Philip, Chacon, and Pernice [26], a reduced set of MHD equations is obtained that simulate a "large aspect-ratio" tokamak with noncircular cross sections. Here the magnetic B-field along the z-direction, or the toroidal direction, is very large and mostly constant.

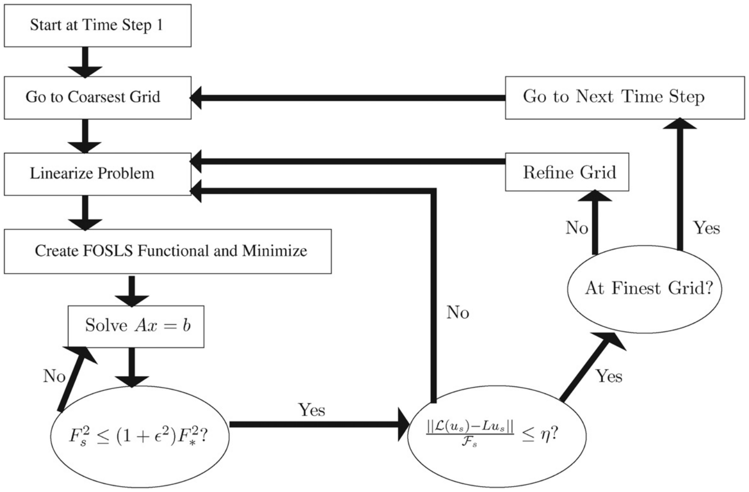

FIG. 3.4. *Flow chart of the nested iteration algorithm. The stopping tolerances* ϵ, *as in* (3.5), *and* η, *as in* (3.15), *are used for the linear iterations and nonlinear iterations, respectively.*

In this context, the plasma behavior in the poloidal cross section is studied. This was described in the companion paper, [3]. In that and related papers [11, 12], convergence of the FOSLS method is also discussed. However, there the nested iteration algorithm was applied on only one time-step. Here, it is applied at each time-step, and we show that the appropriate physics is captured with only a handful of WUs per time-step. In addition, in the companion paper, a time-step convergence analysis is done. This shows that the expected convergence rate for various time-stepping schemes is obtained.

Again the reduced model is equivalent to the 2D version of (2.8)–(2.15). The x-direction denotes the periodic poloidal direction in the tokamak, whereas the y-dimension represents a thin annulus in the poloidal cross section. In this 2D setting, vorticity, ω, and current density, j, are both scalar variables. A biquadratic finite element space and, for the time-stepping, an implicit BDF-2 scheme are used. At all time-steps, the functional is reduced by at least four orders of magnitude; in most cases, the functional value is less than 10^{-3}. All tests were performed on an AMD Opteron shared memory machine running Suse Linux and with 64 GB of shared memory. The software package, FOSPACK [29], was used for all tests. We now apply our methodology to two test problems, known as the tearing mode and the island coalescence problems.

4.1. Test problem: Tearing modes. The first test problem simulates a tearing mode instability arising from perturbations in a steady-state current density sheet of a tokamak generator. In this case, a current density sheet in the toroidal direction of the tokamak is perturbed, resulting in an instability that causes a reconnection in the magnetic field lines. The result is a formation of an island in the contours of the current density field and a "tearing" in the contours of the vorticity. More details can be found in [4, 13, 14, 26]. For the following simulations, we define

$$\Omega = [0,4] \times [0,1], \qquad R_e = S_L = 1000.$$

The initial conditions at equilibrium are

$$\mathbf{B}_0(x,y) = \begin{pmatrix} -\tanh\left(\frac{1}{\lambda}\left(y-\frac{1}{2}\right)\right) + \frac{4}{3\lambda}\left(y-\frac{1}{2}\right)^3 sech^2\left(\frac{\frac{1}{2}}{\lambda}\right) \\ 0 \end{pmatrix}, \tag{4.1}$$

$$\mathbf{u}_0(x,y) = \mathbf{0}, \tag{4.2}$$

$$\omega_0(x,y) = 0, \tag{4.3}$$

$$j_{30}(x,y) = -\frac{4}{\lambda}\left(y-\frac{1}{2}\right)^2 sech^2\left(\frac{\frac{1}{2}}{\lambda}\right) + \frac{1}{\lambda} sech^2\left(\frac{y-\frac{1}{2}}{\lambda}\right), \tag{4.4}$$

$$p_0(x,y) = 0.303 + \frac{1}{2}|\mathbf{B}_0|^2, \tag{4.5}$$

where $\lambda = 0.2$. These initial conditions are perturbed away from equilibrium as follows:

$$\delta\mathbf{B}_0(x,y) = \begin{pmatrix} -\gamma\pi\cos(\pi y)\sin\left(\pi\frac{x}{2}-\frac{\pi}{2}\right) \\ \frac{1}{2}\gamma\pi\sin(\pi y)\cos\left(\pi\frac{x}{2}-\frac{\pi}{2}\right) \\ 0 \end{pmatrix}, \tag{4.6}$$

$$\delta j_{30}(x,y) = -\frac{5}{4}\gamma\pi^2\sin(\pi y)\sin\left(\pi\frac{x}{2}-\frac{\pi}{2}\right), \tag{4.7}$$

$$\delta p_0(x,y) = \frac{5}{8}\gamma^2\pi^2\sin^2(\pi y)\sin^2\left(\pi\frac{x}{2}-\frac{\pi}{2}\right), \tag{4.8}$$

where $\gamma = -0.001$. The boundary conditions are periodic in x and homogeneous Dirichlet for the current density and vorticity on the top and bottom of Ω. Also, $\mathbf{n}\cdot\mathbf{u} = 0$ and $\mathbf{n}\cdot\mathbf{B} = 0$ on the top and bottom. These boundary conditions guarantee that the vorticity FOSLS formulation, (2.8)–(2.15), is H^1 elliptic [2].

4.1.1. Results. The problem was run to time $200\tau_A$ with a time-step of $1\tau_A$ using BDF-2, an implicit time-stepping scheme. Here τ_A is the time in Alfvén units. By the final time-step, the tearing mode has fully developed. The nested iteration algorithm is applied at every time-step, and stopping criteria, as described in section 3, were used for the AMG V-cycles and the linearizations so that each algebraic problem was solved to within a specified fraction of the discretization error. In other words, only a small amount of work was done to obtain the solution on each grid. As the results show, this was sufficient to faithfully produce the relevant physics on a 64×256 fine grid with quadratic elements. This requires $463,239$ degrees of freedom. The linear and nonlinear stopping parameters are

$$F_s^2 \leq 1.01F_*^2, \qquad \frac{F_d}{\mathcal{F}_s} \leq 0.1.$$

On the finest grid and for all time-steps, only one Newton step was needed to approximate the nonlinear solution within the discretization error, and only a small number of V-cycles were needed for the linear solves to achieve the desired accuracy.

Figure 4.1 shows that the method resolves the problem correctly. The main properties of the tearing mode have been captured. The total work is shown in Table 4.1. The WUs for a given time-step are computed by first estimating the work of one V-cycle on a given level, which is defined as the V-cycle complexity, C_v^l, at the given refinement level, l, and computed as follows:

$$C_v^l = \frac{\sum_{j=1}^{m^l} nz_j^l * r_j^l}{nz_1^l}, \tag{4.9}$$

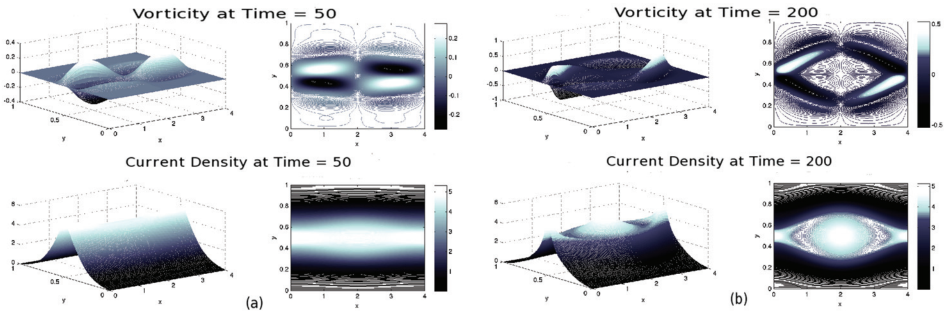

FIG. 4.1. *Numerical solution after (a)* $50\tau_A$ *and (b)* $200\tau_A$. $h = 1/64$ *and* $p = 2$ *using uniform refinement.* $R_e = S_L = 1000$. *Top: vorticity. Bottom: current density.*

where m^l is the total internal levels of a given V-cycle on the overall grid, l, nz_j^l is the number of nonzeros in the linear system on the V-cycle level, and r_j^l is the number of relaxations performed. For all test problems in this paper, V(1,1) cycles are used, and, thus, $r_j^l = 2$. This value is then multiplied by the number of V-cycles performed on that nested grid for all Newton steps. To compare the work from one grid to another, this value is then multiplied by the matrix nonzeros for that grid and summed over all levels. The total work for the time-step is then defined as

$$W_T = \frac{\sum_{l=1}^{n} NZ_l * C_v^l * V_l}{NZ_1}. \tag{4.10}$$

Here $NZ_l = nz_1^l$ is the total nonzeros on each nested iteration grid, NZ_1 refers to the finest grid for that time-step, V_l is the total number of V-cycles performed on the grid, and n is the maximum refinement levels for the given time-step.

The results in Table 4.1 show that, using the nested iteration algorithm, complex instabilities in a fusion reactor simulation are captured at a cost of 30–80 WUs. For later time-steps, a slight deterioration in AMG convergence on finer grids is seen. For these later time-steps, the reconnection is beginning to develop, and, therefore, more features are forming. Thus, the cost of our AMG solver needs to be addressed in future work. It appears that we have scalability, in that the cost is bounded uniformly in the grid size, but that bound is not as small as we expect or would like. Despite this

TABLE 4.1
Number of degrees of freedom, Newton steps, and V-cycles used at each level and time-step. The number of WUs or equivalent fine-grid relaxations are also computed here.

Time	Grid	DOF	Nwt Steps	Avg V-cycles	WUs	Avg WU/Time-step
1-82	6	595	1	4	0.011	≈ 35 WU
1-82	5	2,079	2	11.	0.118	
1-82	4	7,735	2	14.1	0.650	
1-82	3	29,799	2	17.8	3.421	
1-82	2	116,935	1	15.7	12.677	
1-82	1	463,239	1	5.7	18.358	
83-200	6	595	2	8	0.021	≈ 77 WU
83-200	5	2,079	2	11	0.118	
83-200	4	7,735	2	18.4	0.852	
83-200	3	29,799	2	21	4.063	
83-200	2	116,935	1	20.9	16.869	
83-200	1	463,239	1	16.9	55.036	

increased cost, FOSLS gives a discretized system that is symmetric positive definite, and AMG solves it accurately and efficiently, especially in the context of a nested iteration algorithm. It also gives sharp error estimates that aid in continuously calculating the accuracy-per-computational-cost. In addition, as is seen in the proceedings paper, [1], using an efficiency-based AMR scheme reduces the amount of work needed to obtain the same accuracy for the solution.

In the next problem, the island coalescence instability, we again see the effectiveness of the nested iteration algorithm.

4.1.2. Test problem: Island coalescence. The second test problem simulates an island coalescence in the current density arising from perturbations in an initial current density sheet. Again, a current density sheet in the toroidal direction of the tokamak is perturbed, resulting in an instability that causes a reconnection in the magnetic field lines and merging of two islands in the current density field. This produces a sharp peak in current density where the magnetic field lines reconnect. This region is known as the reconnection zone, and the point at which the magnetic field lines break is known as the $\mathcal{X}$-point. See [4, 21, 26] for more details. For the following simulations, we define

$$\Omega = [-1,1] \times [-1,1], \qquad R_e = S_L = 50{,}001.$$

The initial conditions at equilibrium are

$$\mathbf{B}_0(x,y) = \frac{1}{\cosh(2\pi y) + k\cos(2\pi x)} \begin{pmatrix} \sinh(2\pi y) \\ k\sin(2\pi x) \end{pmatrix}, \tag{4.11}$$

$$\mathbf{u}_0(x,y) = \mathbf{0}, \tag{4.12}$$

$$\omega_0(x,y) = 0, \tag{4.13}$$

$$j_{30}(x,y) = \nabla \times \mathbf{B}_0 = \frac{2\pi(k^2-1)}{(\cosh(2\pi y) + 0.2\cos(2\pi x))^2}, \tag{4.14}$$

$$p_0(x,y) = \frac{(1-k^2)}{2}\left(1 + \frac{1}{(\cosh(2\pi y) + 0.2\cos(2\pi x))^2}\right), \tag{4.15}$$

where $k = 0.2$. These initial conditions are perturbed away from equilibrium as follows:

$$\delta\mathbf{B}_0(x,y) = \begin{pmatrix} -\gamma\frac{1}{\pi}\cos(\pi x)\sin\left(\pi\frac{y}{2}\right) \\ \frac{1}{2}\gamma\frac{1}{\pi}\cos\left(\pi\frac{y}{2}\right)\sin(\pi x) \\ 0 \end{pmatrix}, \tag{4.16}$$

$$\delta j_{30}(x,y) = \gamma\cos\left(\pi\frac{y}{2}\right)\cos(\pi x), \tag{4.17}$$

where $\gamma = -0.01$. The boundary conditions are periodic in x and Dirichlet for the current density and vorticity on the top and bottom of the domain. Also $\mathbf{n}\cdot\mathbf{u}$ and $\mathbf{n}\cdot\mathbf{B}$ are known on the top and bottom. Again, the FOSLS formulation, (2.8)–(2.15), is H^1 elliptic.

4.1.3. Results. The problem was run to time $8\tau_A$ with a time-step of $0.1\tau_A$ using a BDF-2 scheme. At the end of the simulation, the islands have begun to coalesce, and a large peak in current density has occurred at the reconnection point. Using uniform refinement and quadratic elements, the nested iteration was performed up to a 128×128 grid. This leads to $462{,}343$ degrees of freedom. The linear and nonlinear stopping parameters are again

$$F_s^2 \le 1.01F_*^2, \qquad \frac{F_d}{\mathcal{F}_s} \le 0.1.$$

As with the tearing mode instability, on the finest grid and for all time-steps, only a couple of Newton steps were needed for the Newton iterations to converge, and a small number of V-cycles were needed for AMG to resolve the linear system. This is seen in Table 4.2. In this problem, when the reconnection begins, a sharp peak in current density occurs, resulting in steep gradients in the magnetic field. As a result, the nonlinearities in the problem get larger for these time-steps. This may explain why, for later time-steps, two Newton steps are needed to get the linear and nonlinear operators to be close on the finest grid. These features might not be seen on the coarser grids, and, therefore, a little more work is needed. On a bigger machine, higher resolutions can be obtained, and we expect the algorithm to eventually need only one Newton step on the finest grids. Future work will simulate the algorithm on a parallel version of the code that will allow for more refinement. However, with the resolution obtained here, the algorithm is still capable of capturing the appropriate physics in only 65 WUs per time-step. The results with uniform refinement at time $8\tau_A$ are seen in Figure 4.2. This shows the sharp peak in the current density, as well as the coalescence of the two islands.

Table 4.2

Number of degrees of freedom, Newton steps, and V-cycles used at each level and time-step. The number of WUs or equivalent fine-grid relaxations are also computed here.

Time	Grid	DOF	Nwt Steps	Avg V-cycles	WUs	Avg WU/Time-step
1-80	7	175	1.2	4.7	0.003	≈ 65 WU
1-80	6	567	1.8	8.7	0.027	
1-80	5	2,023	1	4	0.056	
1-80	4	7,623	1	4.1	0.229	
1-80	3	29,575	1.1	5.7	1.147	
1-80	2	116,487	1.4	11.7	9.532	
1-80	1	462,343	1.5	16.8	54.433	

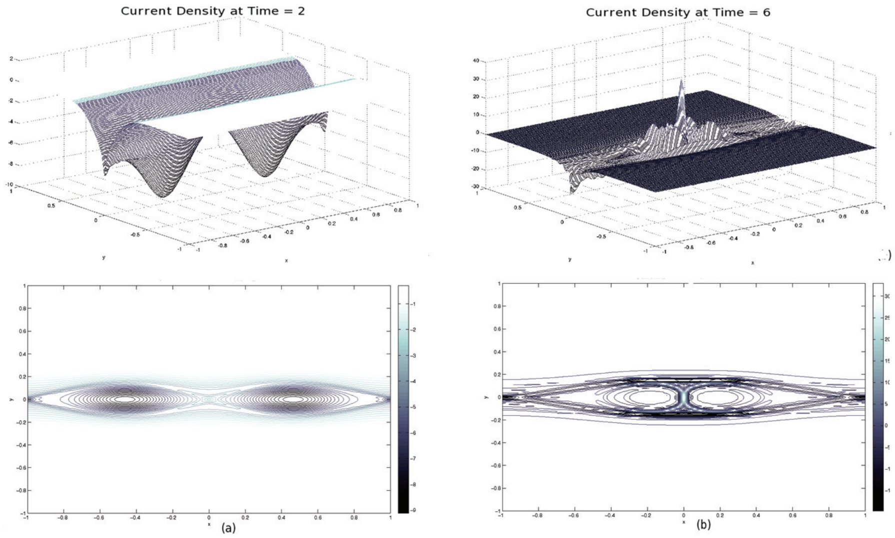

Fig. 4.2. *Numerical solution of current density (j) after (a) $2\tau_A$ and (b) $6\tau_A$. $h = 1/64$ and $p = 2$ using uniform refinement. $R_e = S_L = 50,001$.*

In addition to showing that the algorithm works well numerically on the island problem, we also confirm that it is capable of capturing the physics of the problem. To do this, the growth and reconnection rates of the instabilities are measured. These are described in [4, 13, 21, 26].

First the growth rate, which measures the rate at which the perturbation from the equilibrium state in all the primitive variables grows over time, is computed as follows:

$$\Gamma(t) = log||\mathcal{U} - \mathcal{U}_0||_0. \tag{4.18}$$

Here $\mathcal{U}$ is any of the dependent variables, and $\mathcal{U}_0$ is the unperturbed initial equilibrium state of that variable. We expect the perturbations for the island instability to grow exponentially with time up until the reconnection occurs. This is seen as a linear rate for $\Gamma(t)$. After this, the growth rate should remain steady with order one. This is shown in Figure 4.3 for Reynolds and Lundquist numbers of 5,000 and 50,001. The instabilities are propagating at the expected rate.

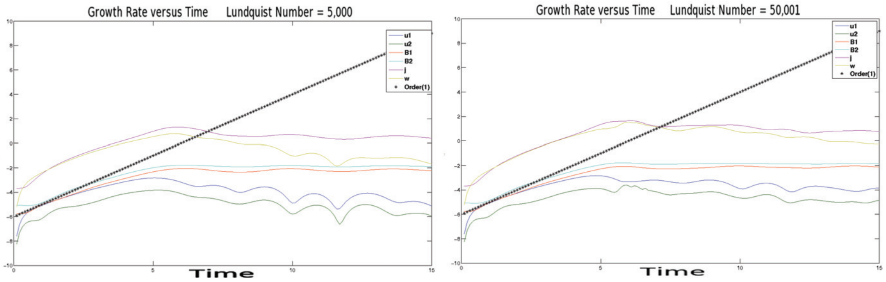

FIG. 4.3. *Growth rate of all variables over time. Left: $R_e = S_L = 5,000$. Right: $R_e = S_L = 50,001$.*

Another measure of the quality of the approximate solution is the reconnection rate. This describes how the magnetic field lines break and reconnect over time. It is measured by checking how the current density, j, changes with time relative to its equilibrium state, proportionally to the resistivity of the system, specifically at the point where the reconnection occurs. For the island coalescence problem, this is known as the $\mathcal{X}$-point or the point at which the two islands collide. We have

$$E_z|_{\mathcal{X}} = \frac{\partial \Psi}{\partial t}|_{\mathcal{X}} = \frac{1}{S_L}(j(t) - j(0))|_{\mathcal{X}}, \tag{4.19}$$

where Ψ is the poloidal flux function. As in [13, 21], it is expected that the reconnection rate increases over time until the reconnection occurs, when it reaches its peak magnitude. It then decays back to zero as the islands merge into one. For higher Lundquist numbers (i.e., lower resistivities), this decay involves some oscillations known as "sloshing" [21]. This behavior is also observed as shown in Figure 4.4. For a Lundquist number of $5,000$, the reconnection rate peaks and then smoothly decays to zero. For Lundquist numbers of $10,000$ and $50,001$, the sloshing develops after the reconnection has occurred. In addition, it is predicted that the peak reconnection rate decreases with Lundquist number with order $O(S_L^{-1/2})$. The nested iteration method captures this relation, as shown in Figure 4.4.

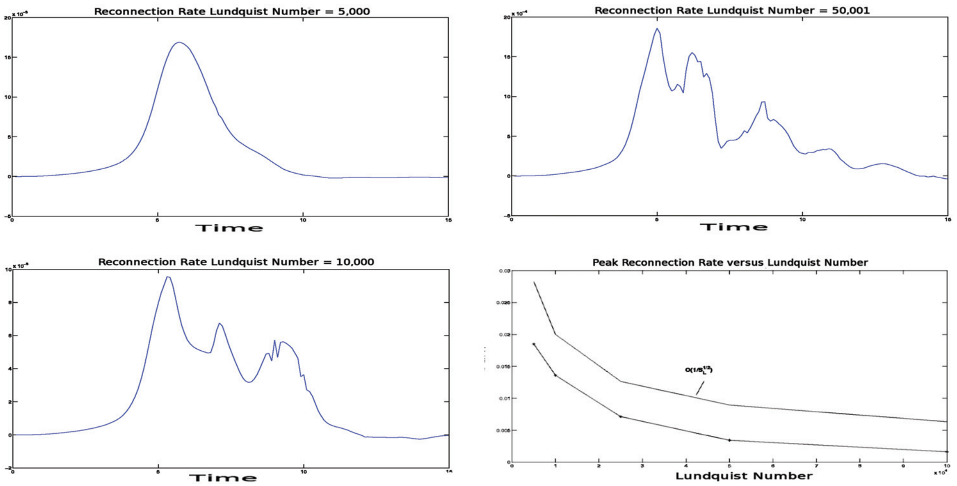

FIG. 4.4. *Top and bottom left: reconnection rate versus time. Top left:* $S_L = 5,000$. *Top right:* $S_L = 50,001$. *Bottom left:* $S_L = 10,000$. *Bottom right: peak reconnection rate versus Lundquist number,* S_L.

5. Discussion. We showed that the FOSLS finite element method, along with nested iteration, is highly effective for the complicated, current-vorticity form of resistive MHD. Real-world MHD applications are solved very efficiently when the focus is on accuracy-per-computational-cost. The crux of the above methodology is that most of the difficult computations, such as the linearizations, are done on coarser grids, where computing is less expensive. Moreover, the result of interpolating the solution from the coarse grid up to the fine grid provides a good starting guess for the solution on that finer grid. As a result, when the desired accuracy is reached, an approximation to the solution is obtained in only a few WUs. The use of FOSLS greatly aids this process. Its sharp, a posteriori error estimate allows parameters to be computed that are used to estimate the current accuracy-per-computational-cost. However, any other discretization method and iterative linear solver that can provide such parameters can be used instead.

While this paper mainly focused on the nested iteration concepts, an additional way to obtain better accuracy-per-computational-cost is through the use of AMR. This allows for the meshes to be locally refined at each time-step and, therefore, follow the physics of the solution as it evolves over time. As stated above, the efficiency-based AMR scheme, ACE [17, 24], has been used successfully on these types of MHD problems [2, 1]. For the island coalescence problem, using the ACE algorithm along with the nested iteration Newton-FOSLS-AMG scheme results in 10% of the work that uniform refinement requires to get to the same functional error.

Several aspects still need to be studied. First, if previous data are used to determine when to move to a finer grid, then the previous data must be solved "well enough" to get good estimates. For instance, is it necessary to solve the problem on the previous time-step to a greater accuracy than otherwise required in order to obtain accurate parameters? For this paper, all of the parameters were chosen based on average estimates and were fixed for the entire run. In the future, we hope to develop heuristics that allow these parameters to be adjusted dynamically. At some point, more complicated physics could occur, and then the algorithm needs to determine those parameters more precisely.

Second, the linear system solvers are what dictate the overall efficiency of the nested iteration Newton-FOSLS method. For this paper, the algebraic systems are solved with a classical algebraic multigrid method. Deteriorations in the algebraic convergence for increased time-step size as well as Reynolds and Lundquist numbers are observed. The current AMG algorithm can be improved in several ways. One might develop an improved AMG scheme for the above type of systems of partial differential equations that might involve the use of newly developed adaptive multigrid algorithms described more in [8, 9, 22]. Using a problem-specific linear solver would produce scalable results and thus would greatly enhance the positive effects of using nested iteration.

Finally, there are many other MHD problems to be tested, as well as other time-dependent problems in fluid dynamics that have large nonlinearities. Using a first-order system least squares formulation with nested iteration, we were able to resolve the above MHD physics, and we believe that, with a careful formulation, these methods can be used for many other time-dependent nonlinear systems.

Acknowledgments. This work was sponsored by the Department of Energy under grants DE-FG02-03ER25574 and DE-FC02-06ER25784, Lawrence Livermore National Laboratory under contract B568677, and the National Science Foundation under grants DMS-0621199, DMS-0749317, and DMS-0811275.

REFERENCES

[1] J. H. Adler, T. A. Manteuffel, S. F. McCormick, J. Nolting, J. W. Ruge, and L. Tang, *An efficiency-based adaptive refinement scheme applied to incompressible, resistive magnetohydrodynamics*, in Large-Scale Scientific Computing, Lecture Notes in Comput. 5910, Sci. Springer, Berlin, 2009, pp. 1–13.

[2] J. H. Adler, *Nested Iteration and First Order Systems Least Squares on Incompressible Resistive Magnetohydrodynamics*, Ph.D. thesis, University of Colorado at Boulder, Boulder, CO, 2009.

[3] J. H. Adler, T. A. Manteuffel, S. F. McCormick, and J. W. Ruge, *First-order system least squares for incompressible resistive magnetohydrodynamics*, SIAM J. Sci. Comput., 32 (2010), pp. 229–248.

[4] G. Bateman, *MHD Instabilities*, The MIT Press, Cambridge, MA, 1978.

[5] A. Brandt, S. F. McCormick, and J. W. Ruge, *Algebraic Multigrid (AMG) for Sparse Matrix Equations*, Cambridge University Press, London, 1984.

[6] A. Brandt, S. McCormick, and J. Ruge, *Algebraic Multigrid (AMG) for Automatic Multigrid Solutions with Application to Geodetic Computations*, report, Institute for Computational Studies, Fort Collins, CO, 1982.

[7] A. Brandt, *Algebraic multigrid theory: the symmetric case*, Appl. Math. Comput., 19 (1986), pp. 23–56.

[8] M. Brezina, R. Falgout, S. MacLachlan, T. Manteuffel, S. McCormick, and J. Ruge, *Adaptive smoothed aggregation (αSA) multigrid*, SIAM Review, 47 (2005), pp. 317–346.

[9] M. Brezina, R. Falgout, S. Maclachlan, T. Manteuffel, S. McCormick, and J. Ruge, *Adaptive algebraic multigrid*, SIAM J. Sci. Comp., 27 (2006), pp. 1261–1286.

[10] W. L. Briggs, V. E. Henson, and S. F. McCormick, *A Multigrid Tutorial*, 2nd ed., Society for Industrial and Applied Mathematics, Philadelphia, PA, 2000.

[11] Z. Cai, R. 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.

[12] Z. Cai, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for second-order partial differential equations: Part II,* SIAM J. Numer. Anal., 34 (1997), pp. 425–454.

[13] L. Chacon, D. A. Knoll, and J. M. Finn, *An implicit, nonlinear reduced resistive MHD solver*, J. Comput. Phys., 178 (2002), pp. 15–36.

[14] L. Chacon, D. A. Knoll, and J. M. Finn, *Nonlinear study of the curvature-driven parallel velocity shear-tearing instability*, Phys. Plasmas, 9 (2002), pp. 1164–1176.

[15] A. L. Codd, T. A. Manteuffel, and S. F. McCormick, *Multilevel first-order system least squares for nonlinear elliptic partial differential equations*, SIAM J. Numer. Anal., 41 (2003), pp. 2197–2209.

[16] A. Codd, *Elasticity-Fluid Coupled Systems and Elliptic Grid Generation (EGG) Based on First-Order System Least Squares (FOSLS)*, Ph.D. thesis, University of Colorado at Boulder, Boulder, CO, 2001.

[17] H. DeSterck, T. Manteuffel, S. McCormick, J. Nolting, J. Ruge, and L. Tang, *Efficiency-based h- and hp-refinement strategies for finite element methods*, Numer. Linear Algebra Appl., 15 (2008), pp. 249–270.

[18] P. Deuflhard, *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms,* Springer, Berlin, 2004.

[19] D. S. Harned and D. D. Schnack, *Semi-implicit method for long time scale magnetohydrodynamic calculations in three dimensions*, J. Comput. Phys., 65 (1986), pp. 57–70.

[20] O. S. Jones, U. Shumlak, and D. S. Eberhardt, *An implicit scheme for nonideal magnetohydrodynamics*, J. Comput. Phys., 130 (1997), pp. 231–242.

[21] D. A. Knoll and L. Chacon, *Coalescence of magnetic islands, sloshing, and the pressure problem*, Phys. Plasmas, 13 (2006), pp. 032307.1–032307.7.

[22] S. MacLachlan, *Improving Robustness in Multiscale Methods*, Ph.D. thesis, University of Colorado at Boulder, Boulder, CO, 2004.

[23] D. R. Nicholson, *Introduction to Plasma Theory*, John Wiley and Sons, New York, 1983.

[24] J. Nolting, *Efficiency-based Local Adaptive Refinement for FOSLS Finite Elements*, Ph.D. thesis, University of Colorado at Boulder, Boulder, CO, 2008.

[25] U. Trottenberg, C. W. Oosterlee, and A. Schuller, *Multigrid*, Academic Press, New York, 2000.

[26] B. Philip, L. Chacon, and M. Pernice, *Implicit adaptive mesh refinement for 2D reduced resistive magnetohydrodynamics*, J. Comput. Phys., 227 (2008), pp. 8855–8874.

[27] O. Roehrle, *Multilevel First-Order System Least Squares for Quasilinear Elliptic Partial Differential Equations*, Ph.D. thesis, University of Colorado at Boulder, Boulder, CO, 2004.

[28] J. Ruge and K. Stüben, *Algebraic Multigrid (AMG)*, in Multigrid Methods, Frontiers Appl. Math. 3, S.F. McCormick, ed., SIAM, Philadelphia, PA, 1987, pp. 73–130.

[29] J. Ruge, Fospack users manual, version 1.0. unpublished, 2000.

[30] G. Starke, *Gauss–Newton multilevel methods for least-squares finite element computations of variably saturated subsurface flow*, Computing, 64 (2000), pp. 323–338.

[31] H. R. Strauss, *Nonlinear, three-dimensional magnetohydrodynamics of noncircular tokamaks*, Phys. Fluids, 19 (1976), pp. 134–140.

[32] P. Ullrich, *Dynamics and Properties of the Magnetohydrodynamics Equations*, unpublished, 2005.

[33] C. Westphal, *First-Order System Least Squares for Geometrically-Nonlinear Elasticity in Nonsmooth Domains*, Ph.D. thesis, University of Colorado at Boulder, Boulder, CO, 2004.