SIAM J. NUMER. ANAL.
Vol. 39, No. 4, pp. 1418–1445

FIRST-ORDER SYSTEM $\mathcal{LL}^*$ (FOSLL*): SCALAR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS *

Z. CAI†, T. A. MANTEUFFEL‡‡, S. F. MCCORMICK‡, AND J. RUGE‡

Abstract. The L^2-norm version of first-order system least squares (FOSLS) attempts to reformulate a given system of partial differential equations so that applying a least-squares principle yields a functional whose bilinear part is H^1-elliptic. This ellipticity means that the minimization process amounts to solving a weakly coupled system of Poisson-like scalar equations. An unfortunate limitation of the L^2-norm FOSLS approach is that this product H^1 equivalence generally requires sufficient smoothness of the original problem. Inverse-norm FOSLS overcomes this limitation, but at a substantial loss of real efficiency. The FOSLL* approach introduced here is a promising alternative that is based on recasting the original problem as a minimization principle involving the adjoint equations. This paper provides a theoretical foundation for the FOSLL* methodology and illustrates its performance by applying it numerically to several examples. Results for the so-called two-stage approach applied to discontinuous coefficient problems show promising robustness and optimality. Indeed, FOSLL* appears to exhibit the generality of the inverse-norm FOSLS approach while retaining the full efficiency of the L^2-norm approach.

Key words. least squares, elliptic, multigrid

AMS subject classifications. 65F10, 65F30

PII. S0036142900388049

1. Introduction.

First-order system least squares (FOSLS) was developed for numerical solution of a wide range of partial differential equations (see [1, 2, 3, 4, 7, 9, 10, 11, 12, 13, 14, 19, 16, 18, 19] and references therein). The basic idea behind the standard FOSLS approach is that it recasts the original system as an expanded first-order system to which a least L^2-norm principle is applied. Its central aim is to reformulate the original system as the minimization of a functional whose bilinear part is equivalent to the product H^1 norm (i.e., the square root of the bilinear part is continuous and coercive in the norm formed by summing the H^1 norms applied to each variable). This product H^1 equivalence means that the minimization process amounts to solving a weakly coupled system of Poisson-like scalar equations. This property in turn implies that standard discretization methods can be used to achieve optimal H^1 accuracy in all variables, with resulting symmetric positive-definite matrix equations that can be efficiently solved by standard multigrid.

One limitation of L^2-norm FOSLS is that product H^1 equivalence usually can be confirmed only under sufficient smoothness assumptions on the original problem (e.g., the domain, coefficients, and data). Inverse-norm versions of FOSLS can overcome this limitation (cf. [4, 9, 12]), but at the expense of rather awkward norm evaluation requirements and the attendant loss of full efficiency. In fact, because the inverse-norm

*Received by the editors April 25, 2000; accepted for publication April 4, 2001; published electronically October 23, 2001.

http://www.siam.org/journals/sinum/39-4/38804.html

†Department of Mathematics, Purdue University, 1395 Mathematical Sciences Building, West Lafayette, IN 47907-1395 (zcai@math.purdue.edu). The work of this author was sponsored by the National Science Foundation under grants DMS-9619792 and INT-9910010.

‡Department of Applied Mathematics, Campus Box 526, University of Colorado at Boulder, Boulder, CO 80309-0526 (tmanteuf@boulder.colorado.edu, stevem@boulder.colorado.edu, jruge@boulder.colorado.edu). The work of these authors was sponsored by the National Science Foundation under grant DMS-9706866 and by the U.S. Department of Energy under grant DE-FG03-93ER25165.

usually does not take into account the underlying problem (e.g., varying coefficients), the constants in the inverse-norm continuity and coercivity bounds can vary widely. Since solution methods are typically dependent on the spread of these constants, costs can be much larger for the inverse-norm than for the L^2-norm FOSLS approach.

Our purpose here is to develop a new approach that also begins by recasting the original problem as an expanded first-order system $\mathcal{L}\mathbf{u} = \mathbf{f}$. However, what we call the FOSLL* approach now departs from FOSLS because it obtains the minimization principle by first rewriting the system as $\mathcal{L}\mathcal{L}^*\mathbf{w} = \mathbf{f}$ in terms of the dual variables $\mathbf{w}$ and the adjoint $\mathcal{L}^*$. The functional to be minimized is then just $\|\mathcal{L}^*\mathbf{w}\|^2 - 2\,\langle \mathbf{w}\,,\mathbf{f}\rangle$.

FOSLL* seems to retain the full efficiency of the L^2-norm FOSLS approach while achieving the generality of the inverse-norm approach. Our theory confirms the generality of the method, and our numerical results illustrate its efficiency. It is especially noteworthy that the so-called two-stage FOSLL* approach (see section 5) applied to discontinuous coefficient problems achieves finite-element accuracy and multigrid efficiency that is typical of standard methods for Poisson-like problems.

The robustness and efficiency of this new approach is demonstrated below by applying FOSLL* to a general elliptic scalar equation. However, for clarity, we first discuss the genesis of the methodology in abstract terms.

The FOSLL* method is basically a dual of the FOSLS approach. We take our cue here from the matrix problem $Ax = b$: Least-squares methods involve matrices of the form A^tA that arise from minimizing $\|Ax - b\|^2$; the dual of this method involves matrices of the form AA^t that arise from knowing that $Ax = b$ has a solution if and only if $AA^ty = b$ does; and $x = A^ty$ is the minimal norm solution of $Ax = b$. Note that solving $Ax = b$ by this dual approach can be done entirely with the x variable, just as Kaczmarz's method can be viewed as a Gauss–Seidel method applied to $AA^ty = b$ but translated to approximations of x [20].

Care must be taken to ensure that the $\mathcal{L}$ we construct in the differential setting has properties that come naturally to the finite-dimensional setting described above. Suppose that $\mathcal{L}$ has been constructed, with domain $\mathcal{D}$ and range $\mathcal{R}$, so that the solution $\mathbf{u}$ we seek satisfies $\mathcal{L}\mathbf{u} = \mathbf{f}$. Further, suppose that $\mathcal{L}^*$ has domain $\mathcal{D}^*$ and range $\mathcal{R}^*$, where $\mathcal{D}^*$ is a Hilbert space, and is continuous and with continuous inverse (see (2.5), (2.6), and section 2 for details of what follows), and suppose that $\mathbf{u}$ is in the range of $\mathcal{L}^*$. Applying FOSLS to $\mathcal{L}\mathbf{u} = \mathbf{f}$ with an inverse norm that takes $\mathcal{L}$ into account leads naturally to the dual norm

$$\|\mathbf{v}\|_B := \sup_{\mathbf{w}\in\mathcal{D}^*} \frac{\langle \mathbf{v},\, \mathbf{w}\rangle}{\|\mathcal{L}^*\mathbf{w}\|} = \langle \mathbf{v},\, \mathbf{z}\rangle^{\frac{1}{2}}\,, \tag{1.1}$$

where $\mathbf{z}$ is the unique element of $\mathcal{D}^*$ such that

$$\langle \mathcal{L}^*\mathbf{z},\, \mathcal{L}^*\mathbf{w}\rangle = \langle \mathbf{v},\, \mathbf{w}\rangle \tag{1.2}$$

for every $\mathbf{w}$ in $\mathcal{D}^*$. Subscript B here refers to the operator relating $\mathbf{v}$ and $\mathbf{z}$: $B\mathbf{v} = \mathbf{z}$. In the finite-dimensional example, $B = (\mathcal{L}\mathcal{L}^*)^{-1}$, which implies $\|\mathbf{v}\|_B = \|(\mathcal{L}\mathcal{L}^*)^{-\frac{1}{2}}\mathbf{v}\|$. The resulting dual-norm FOSLS functional is then given by

$$F(\mathbf{v};\, \mathbf{f}) := \|\mathcal{L}\mathbf{v} - \mathbf{f}\|_B^2\,. \tag{1.3}$$

If both $\mathbf{v}$ and $\mathbf{u}$ are in the range of $\mathcal{L}^*$, then

$$F(\mathbf{v}; \mathbf{f}) := \|\mathcal{L}(\mathbf{v}-\mathbf{u})\|_B^2 = \left(\sup_{\mathbf{w}\in\mathcal{D}^*} \frac{\langle \mathcal{L}(\mathbf{v}-\mathbf{u}),\, \mathbf{w}\rangle}{\|\mathcal{L}^*\mathbf{w}\|} \right)^2$$
$$= \left(\sup_{\mathbf{w}\in\mathcal{D}^*} \frac{\langle \mathbf{v}-\mathbf{u},\, \mathcal{L}^*\mathbf{w}\rangle}{\|\mathcal{L}^*\mathbf{w}\|} \right)^2 = \|\mathbf{v}-\mathbf{u}\|^2.$$

Thus, minimizing $F(\mathbf{v}; \mathbf{f})$ over $\mathbf{v}$ in the range of $\mathcal{L}^*$ is *precisely* the same as minimizing $E(\mathbf{y}; \mathbf{f}) := \|\mathcal{L}^*\mathbf{y} - \mathbf{u}\|$, the L^2-norm of the error, over the *dual variables* $\mathbf{y} \in \mathcal{D}^*$. Minimizing $E(\mathbf{y}; \mathbf{f})$ over $\mathbf{y} \in \mathcal{D}^*$ is accomplished by solving the weak problem of finding $\mathbf{x} \in \mathcal{D}^*$ such that

$$\langle \mathcal{L}^*\mathbf{x},\, \mathcal{L}^*\mathbf{w}\rangle = \langle \mathbf{u},\, \mathcal{L}^*\mathbf{w}\rangle = \langle \mathcal{L}\mathbf{u},\, \mathbf{w}\rangle = \langle \mathbf{f},\, \mathbf{w}\rangle \tag{1.4}$$

for every $\mathbf{w} \in \mathcal{D}^*$. The solution we seek is then $\mathbf{u} = \mathcal{L}^*\mathbf{x}$.

Weak problem (1.4) has a unique solution if $\mathcal{L}^*$ is continuous with continuous inverse (see section 2). Thus, our goal is to design operator $\mathcal{L}$ and problem $\mathcal{L}\mathbf{u} = \mathbf{f}$ so that $\mathcal{L}^*$ has a continuous inverse, $\mathbf{u}$ is in the range of $\mathcal{L}^*$, and weak problem (1.4) is easy to approximate computationally. Note that minimizing $E(\mathbf{y}; \mathbf{f})$ over subspace S is *precisely* the same as minimizing the L^2-norm of the error in the approximation from $\mathcal{L}^*S$.

FOSLL* can be done with all computations taking place in the primitive variable $\mathbf{u}$ instead of the dual variable $\mathbf{x}$. This is generally not an advantage, however, because the natural multigrid coarsening process would reduce only the number of equations, not the number of variables, with the result that the multigrid solver would be much more expensive than the approach we introduce below.

In this paper, FOSLL* is applied to a general scalar elliptic problem. Formulations for other systems of equations, such as Stokes and linear elasticity, will appear in a future study. To this end, let Ω be a bounded, open, connected domain in $\Re^d$ ($d = 2$ or 3) with Lipschitz boundary $\partial\Omega$. Consider the following second-order elliptic boundary value problem:

$$\begin{cases} -\nabla\cdot A\nabla p + \mathbf{b}\cdot\nabla p + cp &= f \quad \text{in} \quad \Omega, \\ p &= 0 \quad \text{on} \quad \Gamma_D, \\ \mathbf{n}\cdot A\nabla p &= 0 \quad \text{on} \quad \Gamma_N, \end{cases} \tag{1.5}$$

where the symbols $\nabla\cdot$ and ∇ stand for the divergence and gradient operators, respectively, A is a $d\times d$ symmetric matrix of functions in $L^\infty(\Omega)$, $\mathbf{b}$ is a $d\times 1$ vector of functions in $L^2(\Omega)$ such that $\nabla\cdot\mathbf{b} \in L^\infty(\Omega)$, c is a real-valued function in $L^\infty(\Omega)$, $\Gamma_D \cup \Gamma_N = \Gamma$ is a partition of the boundary of Ω, and $\mathbf{n} = (n_1, \dots, n_d)^t$ is the outward unit vector normal to the boundary. For simplicity, we assume that both Γ_D and Γ_N are nonempty, with the obvious generalization to quotient spaces when one of them is empty. Assume that A is uniformly elliptic: There exist positive constants λ and Λ such that

$$\lambda\boldsymbol{\xi}^t\boldsymbol{\xi} \le \boldsymbol{\xi}^t A\boldsymbol{\xi} \le \Lambda\boldsymbol{\xi}^t\boldsymbol{\xi}$$

for all $\boldsymbol{\xi} \in R^d$ and almost all $x \in \overline{\Omega}$.

We assume that (1.5) has no nontrivial solution when $f = 0$ and that the adjoint boundary value problem

$$\begin{cases} -\nabla\cdot A\nabla p - \nabla\cdot(\mathbf{b}p) + cp &= f \quad \text{in} \quad \Omega, \\ p &= 0 \quad \text{on} \quad \Gamma_D, \\ \mathbf{n}\cdot(A\nabla p + \mathbf{b}p) &= 0 \quad \text{on} \quad \Gamma_N \end{cases} \tag{1.6}$$

also has no nontrivial solution when $f = 0$.

There are many ways to develop FOSLL* methods for solving (1.5). In the next section, we recall some general results that help us construct the first-order operator $\mathcal{L}$ with the desired properties. In section 3, we introduce notation, recall the L^2-norm version of FOSLS, and show that the resulting $\mathcal{L}$ is not suitable for FOSLL*. In section 4, we develop an approach based on extending $\mathcal{L}$ to more dependent variables so that both $\mathcal{L}$ and $\mathcal{L}^*$ are bijective (1-to-1 and onto) and continuous with continuous inverses on $\mathcal{D}$ and $\mathcal{D}^*$, respectively. We introduce an approach called a two-stage method, which is based on a combination of FOSLS and FOSLL*, in section 5. This is especially useful in that all computation can be done in standard discrete subspaces of H^1 even in the presence of discontinuities in the coefficient matrix A. Discretization error estimates for both methods are obtained in section 6. We conclude with numerical experiments in section 7. One important observation is that, for problems with large convection, FOSLL* efficiency is not limited by the size of the coarsest grid as are many other solvers.

2. Linear operators and adjoints. In this section, we recall some basic definitions about linear operators and their adjoints. Let $\mathcal{V}_1$ and $\mathcal{V}_2$ be Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_{\mathcal{V}_i}$ and norms $\| \cdot \|_{\mathcal{V}_i}$, for $i = 1, 2$, and let $\mathcal{L}$ be a linear operator with domain $\mathcal{D} \subseteq \mathcal{V}_1$ and range $\mathcal{R} \subseteq \mathcal{V}_2$.

If $\mathcal{D}$ is dense in $\mathcal{V}_1$, then we can define the adjoint of $\mathcal{L}$ as follows: If the pair $\mathbf{y} \in \mathcal{V}_2$, $\mathbf{f} \in \mathcal{V}_1$ satisfies

$$\langle \mathcal{L}\mathbf{x}, \mathbf{y} \rangle_{\mathcal{V}_2} = \langle \mathbf{x}, \mathbf{f} \rangle_{\mathcal{V}_1} \qquad \forall \mathbf{x} \in \mathcal{D}, \tag{2.1}$$

then we say $\mathbf{y} \in \mathcal{D}^*$, the domain of $\mathcal{L}^*$, and $\mathbf{f} \in \mathcal{R}^*$, the range of $\mathcal{L}^*$, and write $\mathcal{L}^*\mathbf{y} = \mathbf{f}$.

If $\mathcal{D}^*$ is also dense in $\mathcal{V}_2$, then we may likewise define $(\mathcal{L}^*)^*$ with domain $\mathcal{D}^{**}$. From (2.1), we see that $\mathcal{D} \subseteq \mathcal{D}^{**}$ and $(\mathcal{L}^*)^*\mathbf{x} = \mathcal{L}\mathbf{x}$ $\forall \mathbf{x} \in \mathcal{D}$. The operators we construct later satisfy the property $\mathcal{D} = \mathcal{D}^{**}$, which implies that $(\mathcal{L}^*)^* = \mathcal{L}$, and so we assume it for the remainder of this section.

Suppose that both $\mathcal{D}$ and $\mathcal{D}^*$ are Hilbert spaces under the norms $\| \cdot \|_{\mathcal{D}}$ and $\| \cdot \|_{\mathcal{D}^*}$, respectively, and that $\mathcal{D}$ and $\mathcal{D}^*$ are continuously embedded in $\mathcal{V}_1$ and $\mathcal{V}_2$, respectively: There exist constants c_1 and c_2 such that

$$\|\mathbf{x}\|_{\mathcal{V}_1} \le c_1 \|\mathbf{x}\|_{\mathcal{D}} \qquad \forall \mathbf{x} \in \mathcal{D}, \tag{2.2}$$

$$\|\mathbf{x}\|_{\mathcal{V}_2} \le c_2 \|\mathbf{x}\|_{\mathcal{D}^*} \qquad \forall \mathbf{x} \in \mathcal{D}^*. \tag{2.3}$$

The linear operators $\mathcal{L}$ and $\mathcal{L}^*$ developed in this paper can be viewed from two perspectives. First, they are mutually adjoint linear operators on the Hilbert spaces $\mathcal{V}_1$ and $\mathcal{V}_2$. Second, they are linear operators from Hilbert spaces $\mathcal{D}$ and $\mathcal{D}^*$ to the Hilbert spaces $\mathcal{V}_2$ and $\mathcal{V}_1$, respectively. We use properties of both of these roles to obtain our results.

Toward this end, assume that $\mathcal{L}$ is continuous from $\mathcal{D}$ into $\mathcal{V}_2$ and $\mathcal{L}^*$ is continuous from $\mathcal{D}^*$ into $\mathcal{V}_1$: There exist constants c_3 and c_4 such that

$$\|\mathcal{L}\mathbf{x}\|_{\mathcal{V}_2} \le c_3 \|\mathbf{x}\|_{\mathcal{D}} \qquad \forall \mathbf{x} \in \mathcal{D}, \tag{2.4}$$

$$\|\mathcal{L}^*\mathbf{x}\|_{\mathcal{V}_1} \le c_4 \|\mathbf{x}\|_{\mathcal{D}^*} \qquad \forall \mathbf{x} \in \mathcal{D}^*. \tag{2.5}$$

If $\mathcal{L}$ is injective (1-to-1), then we may consider $\mathcal{L}^{-1}$ as an operator from $\mathcal{R} \subseteq \mathcal{V}_2$ into Hilbert space $\mathcal{D}$. In this context, $\mathcal{L}^{-1}$ is continuous from $\mathcal{R}$ to $\mathcal{D}$ if there exists constant c_0 such that

$$c_0\|\mathbf{x}\|_{\mathcal{D}} \leq \|\mathcal{L}\mathbf{x}\|_{\mathcal{V}_2} \qquad \forall \mathbf{x} \in \mathcal{D}. \tag{2.6}$$

Equivalently, we say that the bilinear form

$$a(\mathbf{x}, \mathbf{w}) := \langle \mathcal{L}\mathbf{x}, \mathcal{L}\mathbf{w} \rangle_{\mathcal{V}_2} \tag{2.7}$$

is coercive on $\mathcal{D}$.

If we assume that $(\mathcal{L}^*)^{-1}$ is continuous from $\mathcal{R}^*$ to $\mathcal{D}^*$, then the bilinear form

$$a^*(\mathbf{x}, \mathbf{w}) := \langle \mathcal{L}^*\mathbf{x}, \mathcal{L}^*\mathbf{w} \rangle_{\mathcal{V}_1} \tag{2.8}$$

is coercive on $\mathcal{D}^*$, and the Lax–Milgram lemma [15] implies that, for every $\mathbf{f} \in \mathcal{V}_2$, there is a unique $\mathbf{x} \in \mathcal{D}^*$ that satisfies

$$\langle \mathcal{L}^*\mathbf{x}, \mathcal{L}^*\mathbf{w} \rangle_{\mathcal{V}_1} = \langle \mathbf{f}, \mathbf{w} \rangle_{\mathcal{V}_2} \qquad \forall \mathbf{w} \in \mathcal{D}^*. \tag{2.9}$$

The above assumptions also imply the following simple results:

1. $\mathcal{R}^*$ is closed in $\mathcal{V}_1$.
2. If $\mathcal{R}^* \neq \mathcal{V}_1$, then $(\mathcal{R}^*)^\perp = \mathcal{N}$, the null space of $\mathcal{L}$.
3. If $\mathcal{R}^* = \mathcal{V}_1$, then
   i. $(\mathcal{L}^*)^{-1} : \mathcal{V}_1 \to \mathcal{D}^*$ is well defined and bounded;
   ii. $\mathcal{L}$ is injective;
   iii. $\mathcal{L}^{-1} : \mathcal{R} \to \mathcal{D}$ is well defined and bounded;
   iv. $(\mathcal{L}^*)^{-1} = (\mathcal{L}^{-1})^*$ $\;(:= \mathcal{L}^{-*})$;
   v. $\mathcal{R} = \mathcal{V}_2$;
   vi. $\mathcal{L}^{-1}$ is continuous from $\mathcal{R} \subseteq \mathcal{V}_2$ to $\mathcal{D}$.

We reword the above results in a manner that will be useful later.

LEMMA 2.1. *Suppose that $\mathcal{L} : \mathcal{D} \subseteq \mathcal{V}_1 \to \mathcal{V}_2$ is a linear operator from Hilbert space $\mathcal{D}$ into Hilbert space $\mathcal{V}_2$ and that $\mathcal{D}$ is continuously embedded and dense in $\mathcal{V}_1$. Let $\mathcal{L}^* : \mathcal{D}^* \subseteq \mathcal{V}_2 \to \mathcal{V}_1$ be its adjoint and suppose that $\mathcal{D}^*$ is also a Hilbert space that is continuously embedded and dense in $\mathcal{V}_2$. Further, suppose that $(\mathcal{L}^*)^* = \mathcal{L}$ and that $\mathcal{L}$ and $\mathcal{L}^*$ are continuous from $\mathcal{D}$ and $\mathcal{D}^*$ into $\mathcal{V}_1$ and $\mathcal{V}_2$, respectively* (*i.e., they satisfy* (2.4) *and* (2.5)). *Then $\mathcal{L}$ is bijective* (1*-to-*1 *and onto*) *and has continuous inverse* (*i.e., satisfies* (2.6)) *if and only if $\mathcal{L}^*$ is bijective and has continuous inverse.*

Proof. We offer only a brief proof. If $\mathcal{L}$ is bijective with continuous inverse, then, using (2.2) and (2.6), we see that $\mathcal{L}^{-1} : \mathcal{V}_2 \to \mathcal{D}$ is a bounded linear operator on $\mathcal{V}_2$. Thus, $\mathcal{L}^*$ is also injective (1-to-1) and $\mathcal{L}^{-*} := (\mathcal{L}^*)^{-1} = (\mathcal{L}^{-1})^* : \mathcal{R}^* \to \mathcal{D}^*$ is also bounded with norm $\|\mathcal{L}^{-*}\| = \|\mathcal{L}^{-1}\|$. This implies that $\mathcal{R}^*$ is closed in $\mathcal{V}_1$. Finally, $\mathcal{R}^* = \mathcal{V}_1$ because $\mathcal{L}$ has only the trivial null space.

We have now established that $\mathcal{L}^*$ is a continuous bijective map from $\mathcal{D}^*$ to $\mathcal{V}_1$. The closed graph theorem implies that $\mathcal{L}^{-*}$ is also continuous. The converse follows from a similar argument. This completes the proof. ∎

LEMMA 2.2. *Under the same assumptions as in Lemma* 2.1, *if both $\mathcal{L}^{-1}$ and $\mathcal{L}^{-*}$ are continuous, then $\mathcal{R} = \mathcal{V}_2$ and $\mathcal{R}^* = \mathcal{V}_1$.*

Proof. If $\mathcal{L}^{-1}$ and $\mathcal{L}^{-*}$ are continuous, then both $\mathcal{L}$ and $\mathcal{L}^*$ are injective and both $\mathcal{R}$ and $\mathcal{R}^*$ are closed. If $\mathcal{R} \neq \mathcal{V}_2$, then there exists some $\mathbf{z} \in \mathcal{V}_2$ such that $\mathbf{z} \perp \mathcal{R}$, which implies that $\mathbf{z} \in \mathcal{N}^*$, the null space of $\mathcal{L}^*$, which in turn contradicts the assumption that $\mathcal{L}^*$ is injective. A similar argument yields $\mathcal{R}^* = \mathcal{V}_1$, which completes the proof. ∎

In the next section, we introduce notation and demonstrate why standard FOSLS formulations must be modified to exploit the FOSLL* strategy. In section 4, we construct $\mathcal{L}$ such that both $\mathcal{L}$ and $\mathcal{L}^*$ are bijective and both $\mathcal{L}^{-1}$ and $\mathcal{L}^{-*}$ are continuous. Thus, the solution of the following weak problem automatically yields $\mathbf{u} = \mathcal{L}^*\mathbf{x}$ as the unique solution of $\mathcal{L}\mathbf{u} = \mathbf{f}$: Find $\mathbf{x} \in \mathcal{D}^*$ that satisfies

$$\langle \mathcal{L}^*\mathbf{x},\, \mathcal{L}^*\mathbf{w}\rangle_{\mathcal{V}_1} = \langle \mathbf{f},\, \mathbf{w}\rangle_{\mathcal{V}_2} \qquad \forall \mathbf{w} \in \mathcal{D}^*. \tag{2.10}$$

To see this, note that (2.10) yields

$$\langle \mathcal{L}^*\mathbf{x},\, \mathcal{L}^*\mathbf{w}\rangle_{\mathcal{V}_1} = \langle \mathbf{f},\, \mathbf{w}\rangle_{\mathcal{V}_2} = \langle \mathbf{u},\, \mathcal{L}^*\mathbf{w}\rangle_{\mathcal{V}_1} \qquad \forall \mathbf{w} \in \mathcal{D}^*.$$

Since $\mathcal{R}^* = \mathcal{V}_1$, this yields $\mathbf{u} = \mathcal{L}^*\mathbf{x}$.

3. Motivation and notation. In this section, we recall the basic FOSLS system for boundary value problem (1.5), develop notation that we need in the remainder of the paper, and motivate the development of the extended system described in the next section.

We primarily use notation for the case $d = 3$ and consider the special case $d = 2$ in the natural way by identifying $\Re^2$ with the $(x_1,\, x_2)$-plane in $\Re^3$. Thus, if $\mathbf{u}$ is two-dimensional, then the curl of $\mathbf{u} = (u_1,\, u_2)^t$ means the scalar function

$$\nabla\times \mathbf{u} = \partial_1 u_2 - \partial_2 u_1$$

and $\nabla^\perp$ denotes its formal adjoint defined by

$$\nabla^\perp q = \begin{pmatrix} \partial_2 q \\ -\partial_1 q \end{pmatrix}.$$

Let $\mathbf{u}$ be a new vector variable satisfying

$$\mathbf{u} = A^{\frac{1}{2}}\nabla p.$$

Equation (1.5) can then be written

$$-\nabla\cdot A^{\frac{1}{2}}\mathbf{u} + \mathbf{b}\cdot A^{-\frac{1}{2}}\mathbf{u} + cp = f.$$

Notice that $\mathbf{u}$ satisfies the constraint

$$\nabla\times A^{-\frac{1}{2}}\mathbf{u} = \mathbf{0}$$

and the boundary condition

$$\mathbf{n}\times A^{-\frac{1}{2}}\mathbf{u} = 0 \quad \text{on} \quad \Gamma_D.$$

Putting these equations together yields the first-order system

$$\left\{\begin{array}{rcll} \mathbf{u} - A^{\frac{1}{2}}\nabla p &=& \mathbf{0} & \text{in} \quad \Omega, \\ -\nabla\cdot A^{\frac{1}{2}}\mathbf{u} + \mathbf{b}\cdot A^{-\frac{1}{2}}\mathbf{u} + cp &=& f & \text{in} \quad \Omega, \\ \nabla\times A^{-\frac{1}{2}}\mathbf{u} &=& \mathbf{0} & \text{in} \quad \Omega, \end{array}\right. \tag{3.1}$$

which we denote as $\mathcal{L}_b(\mathbf{u},\, p)^t = (\mathbf{0}, f, \mathbf{0})$, with boundary conditions

$$\left\{\begin{array}{rcll} \mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{u} &=& 0 & \text{on} \quad \Gamma_N, \\ \mathbf{n}\times A^{-\frac{1}{2}}\mathbf{u} &=& \mathbf{0} & \text{on} \quad \Gamma_D, \\ p &=& 0 & \text{on} \quad \Gamma_D. \end{array}\right. \tag{3.2}$$

To describe the domain and range of $\mathcal{L}_b$, we define spaces that we use in the remainder of the paper. Let

$$H(\operatorname{div} A^{\frac{1}{2}};\,\Omega) := \left\{\mathbf{v} \in L^2(\Omega)^d \,:\, \nabla\cdot A^{\frac{1}{2}}\mathbf{v} \in L^2(\Omega)\right\}, \tag{3.3}$$

$$H(\mathbf{curl}\, A^{-\frac{1}{2}};\,\Omega) := \left\{\mathbf{v} \in L^2(\Omega)^d \,:\, \nabla\times A^{-\frac{1}{2}}\mathbf{v} \in L^2(\Omega)^{2d-3}\right\}, \tag{3.4}$$

which are Hilbert spaces under the respective norms

$$\|\mathbf{v}\|_{H(\operatorname{div} A^{\frac{1}{2}};\,\Omega)} := \left(\|\mathbf{v}\|^2 + \|\nabla\cdot A^{\frac{1}{2}}\mathbf{v}\|^2\right)^{\frac{1}{2}},$$
$$\|\mathbf{v}\|_{H(\mathbf{curl}\, A^{-\frac{1}{2}};\,\Omega)} := \left(\|\mathbf{v}\|^2 + \|\nabla\times A^{-\frac{1}{2}}\mathbf{v}\|^2\right)^{\frac{1}{2}}.$$

When A is the identity matrix, we use the simpler notation $H(\operatorname{div};\,\Omega)$ and $H(\mathbf{curl};\,\Omega)$. Define the subspaces

$$H_J(\operatorname{div} A^{\frac{1}{2}};\,\Omega) := \left\{\mathbf{v} \in H(\operatorname{div} A^{\frac{1}{2}};\,\Omega) \,:\, \mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{v} = 0 \text{ on } \Gamma_J\right\},$$
$$H_J(\mathbf{curl}\, A^{-\frac{1}{2}};\,\Omega) := \left\{\mathbf{v} \in H(\mathbf{curl}\, A^{-\frac{1}{2}};\,\Omega) \,:\, \mathbf{n}\times A^{-\frac{1}{2}}\mathbf{v} = \mathbf{0} \text{ on } \Gamma_J\right\}$$

for $J = D$ and N. Later, we use the spaces

$$\mathcal{W}_{ND}(A) := H_N(\operatorname{div} A^{\frac{1}{2}};\,\Omega) \cap H_D(\mathbf{curl}\, A^{-\frac{1}{2}};\,\Omega), \tag{3.5}$$

$$\mathcal{W}_{DN}(A) := H_D(\operatorname{div} A^{\frac{1}{2}};\,\Omega) \cap H_N(\mathbf{curl}\, A^{-\frac{1}{2}};\,\Omega), \tag{3.6}$$

which are Hilbert spaces under the norm

$$\|\mathbf{v}\|_{\mathcal{W}(A)} := \left(\|\mathbf{v}\|^2 + \|\nabla\cdot A^{\frac{1}{2}}\mathbf{v}\|^2 + \|\nabla\times A^{-\frac{1}{2}}\mathbf{v}\|^2\right)^{\frac{1}{2}}. \tag{3.7}$$

When $A = I$, we use the simpler notation $\mathcal{W}_{ND}$ and $\mathcal{W}_{DN}$.

Next, define the following subspaces of $H^1(\Omega)$:

$$H^1_J(\Omega) := \left\{p \in H^1(\Omega) \,:\, p = 0 \text{ on } \Gamma_J\right\} \tag{3.8}$$

for $J = D$ and N. Below, we use $H^{-1}_J(\Omega)$ to denote the dual of $H^1_J(\Omega)$ with norm defined by

$$\|\phi\|_{-1,J} = \sup_{0\neq\psi\in H^1_J(\Omega)} \frac{(\phi,\psi)}{|\psi|_1} \tag{3.9}$$

for $J = D$ and N.

In this context, we have $\mathcal{V}_1 = L^2(\Omega)^{d+1}$ and $\mathcal{V}_2 = L^2(\Omega)^{3d-2}$, and also

$$\mathcal{D} = \mathcal{W}_{ND}(A) \times H^1_D(\Omega), \tag{3.10}$$

which is a Hilbert space under the product norm

$$\|(\mathbf{u}, p)\|_{\mathcal{D}} = \left(\|\mathbf{u}\|^2_{\mathcal{W}(A)} + \|p\|^2_1\right)^{\frac{1}{2}}.$$

System (3.1), (3.2) was studied in [11], where it was shown that $\mathcal{L}_b$ is continuous on $\mathcal{D}$ with continuous inverse.

In this paper, we want to consider $\mathcal{L}_b^*$. If we write $\mathcal{L}_b$ in matrix form,

$$\mathcal{L}_b = \begin{pmatrix} I & -A^{\frac{1}{2}}\nabla \\ \nabla \cdot A^{\frac{1}{2}} - \mathbf{b}\cdot A^{-\frac{1}{2}} & -cI \\ \nabla\times A^{-\frac{1}{2}} & \mathbf{0} \end{pmatrix}, \tag{3.11}$$

then, formally, we have

$$\mathcal{L}_b^* = \begin{pmatrix} I & -A^{\frac{1}{2}}\nabla - A^{-\frac{1}{2}}\mathbf{b} & A^{-\frac{1}{2}}\nabla\times \\ \nabla \cdot A^{\frac{1}{2}} & -cI & \mathbf{0} \end{pmatrix}. \tag{3.12}$$

The domain of $\mathcal{L}_b^*$ involves *dual variables* $(\mathbf{w}, r, \mathbf{x})^t$, with $\mathbf{w} \in H(\operatorname{div} A^{\frac{1}{2}}; \Omega)$, $r \in H^1(\Omega)$, and $\mathbf{x} \in H(\mathbf{curl}; \Omega)$.

Clearly, $\mathcal{L}_b^*$ corresponds to an underdetermined system, so it must be singular and cannot have a continuous inverse. Now, $\mathcal{L}_b$ was designed to be nonsingular so that it could be used with FOSLS. It is, therefore, not surprising that $\mathcal{L}_b^*$ is not so suitable for FOSLL*. What we need here is a way to modify the original equations so that the *adjoint* of the new system operator has a continuous inverse. We do this by adding slack variables, which is analogous to adding equations for FOSLL*. However, instead of starting with (1.5), we find it more convenient to start with (3.1): By carefully adding slack variables to system (3.1), we are able to develop a square nonsingular system whose solution we seek is in the range of the adjoint. This approach is the topic of the next section.

4. FOSLL$_e^*$ for the extended system. In this section, we extend the system for $\mathcal{L}_b$ by adding slack variables to make it square. We show that the extended operator and its adjoint have continuous inverses. Two and three dimensions are treated separately.

4.1. Two dimensions. Consider the following extended first-order system:

$$\left\{ \begin{array}{rcll} \mathbf{u} - A^{\frac{1}{2}}\nabla p + A^{-\frac{1}{2}}\nabla^{\perp} q & = & \mathbf{0} & \text{in } \Omega, \\ -\nabla \cdot A^{\frac{1}{2}}\mathbf{u} + \mathbf{b}\cdot A^{-\frac{1}{2}}\mathbf{u} + cp & = & f & \text{in } \Omega, \\ \nabla\times A^{-\frac{1}{2}}\mathbf{u} & = & 0 & \text{in } \Omega \end{array} \right. \tag{4.1}$$

with boundary conditions

$$\left\{ \begin{array}{rcll} \mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{u} & = & 0 & \text{on } \Gamma_N, \\ \boldsymbol{\tau}\cdot A^{-\frac{1}{2}}\mathbf{u} & = & 0 & \text{on } \Gamma_D, \\ p & = & 0 & \text{on } \Gamma_D, \\ q & = & 0 & \text{on } \Gamma_N. \end{array} \right. \tag{4.2}$$

Here, $\boldsymbol{\tau} = (-n_2, n_1)^t$ denotes the unit tangent oriented counterclockwise on Γ_D. Changing the sign of the second equation in (4.1) for convenience, then the associated differential operator is

$$\mathcal{L}_2 = \begin{pmatrix} I & -A^{\frac{1}{2}}\nabla & A^{-\frac{1}{2}}\nabla^{\perp} \\ \nabla \cdot A^{\frac{1}{2}} - \mathbf{b}\cdot A^{-\frac{1}{2}} & -cI & 0 \\ \nabla\times A^{-\frac{1}{2}} & 0 & 0 \end{pmatrix}, \tag{4.3}$$

and its formal adjoint is

$$\mathcal{L}_2^* = \begin{pmatrix} I & -A^{\frac{1}{2}}\nabla - A^{-\frac{1}{2}}\mathbf{b} & A^{-\frac{1}{2}}\nabla^{\perp} \\ \nabla\cdot A^{\frac{1}{2}} & -cI & 0 \\ \nabla\times A^{-\frac{1}{2}} & 0 & 0 \end{pmatrix}. \tag{4.4}$$

In this context, $\mathcal{L}_2 : \mathcal{D} \subseteq \mathcal{V}_1 \to \mathcal{V}_2$, where $\mathcal{V}_1 = \mathcal{V}_2 = \mathcal{V} := L^2(\Omega)^4$, and $\mathcal{D} = \mathcal{W}_{ND}(A) \times H_D^1 \times H_N^1$, which is a Hilbert space under the product norm

$$\|(\mathbf{u}, p, q)\|_{\mathcal{D}} := \left(\|\mathbf{u}\|^2_{\mathcal{W}(A)} + \|p\|_1^2 + \|q\|_1^2\right)^{\frac{1}{2}} \tag{4.5}$$

and which is compactly embedded in $\mathcal{V}$.

The domain of $\mathcal{L}_2^*$ involves the *dual variables* $(\mathbf{w}, r, s)^t$ with $\mathbf{w} \in H(\operatorname{div} A^{\frac{1}{2}}; \Omega) \cap H(\mathbf{curl}\, A^{-\frac{1}{2}}; \Omega)$ and $r,\, s \in H^1(\Omega)$. To determine the boundary conditions associated with $\mathcal{L}_2^*$, we compute

$$\begin{aligned}
&\langle \mathcal{L}_2\,(\mathbf{u},\, p,\, q)^t,\, (\mathbf{w},\, r,\, s)^t \rangle \\
&\quad = \langle \mathbf{u} - A^{\frac{1}{2}}\nabla p + A^{-\frac{1}{2}}\nabla^{\perp} q,\, \mathbf{w}\rangle + \langle \nabla\cdot A^{\frac{1}{2}}\mathbf{u} - \mathbf{b}\cdot A^{-\frac{1}{2}}\mathbf{u} - cp,\, r\rangle + \langle \nabla\times A^{-\frac{1}{2}}\mathbf{u},\, s\rangle \\
&\quad = \langle \mathbf{u},\, \mathbf{w}\rangle + \langle p,\, \nabla\cdot A^{\frac{1}{2}}\mathbf{w}\rangle + \langle q,\, \nabla\times A^{-\frac{1}{2}}\mathbf{w}\rangle - \langle \mathbf{u},\, A^{\frac{1}{2}}\nabla r\rangle - \langle \mathbf{u},\, A^{-\frac{1}{2}}\mathbf{b} r\rangle - \langle p,\, cr\rangle \\
&\qquad + \langle \mathbf{u},\, A^{-\frac{1}{2}}\nabla^{\perp} s\rangle - \int_{\partial\Omega} p(\mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{w}) - \int_{\partial\Omega} q(\boldsymbol{\tau}\cdot A^{-\frac{1}{2}}\mathbf{w}) \\
&\qquad + \int_{\partial\Omega} (\mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{u}) r + \int_{\partial\Omega} (\boldsymbol{\tau}\cdot A^{-\frac{1}{2}}\mathbf{u}) s \\
&\quad = \langle (\mathbf{u},\, p,\, q)^t,\, \mathcal{L}_2^*\,(\mathbf{w},\, r,\, s)^t\rangle \\
&\qquad - \int_{\Gamma_N} p(\mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{w}) - \int_{\Gamma_D} q(\boldsymbol{\tau}\cdot A^{-\frac{1}{2}}\mathbf{w}) + \int_{\Gamma_D} (\mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{u}) r + \int_{\Gamma_N} (\boldsymbol{\tau}\cdot A^{-\frac{1}{2}}\mathbf{u}) s,
\end{aligned}$$

where (4.2) was used to reduce the boundary integral terms. The remaining boundary integrals vanish for every $(\mathbf{u}, p, q)^t \in \mathcal{D}$ if and only if we enforce the boundary conditions

$$\left\{ \begin{array}{rcll} \mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{w} & = & 0 & \text{on} \quad \Gamma_N, \\ \boldsymbol{\tau}\cdot A^{-\frac{1}{2}}\mathbf{w} & = & 0 & \text{on} \quad \Gamma_D, \\ r & = & 0 & \text{on} \quad \Gamma_D, \\ s & = & 0 & \text{on} \quad \Gamma_N. \end{array} \right. \tag{4.6}$$

Comparing (4.6) with (4.2), we see that $\mathcal{D}^* = \mathcal{D}$.

We next define the bilinear forms

$$\begin{array}{rcl} a_2(\mathbf{u}, p, q;\, \mathbf{v}, \xi, \eta) & := & \langle \mathcal{L}_2(\mathbf{u}, p, q)^t,\, \mathcal{L}_2(\mathbf{v}, \xi, \eta)^t\rangle \\ & = & \langle \mathbf{u} - A^{\frac{1}{2}}\nabla p + A^{-\frac{1}{2}}\nabla^{\perp} q,\, \mathbf{v} - A^{\frac{1}{2}}\nabla\xi + A^{-\frac{1}{2}}\nabla^{\perp}\eta\rangle \\ & + & \langle \nabla\cdot A^{\frac{1}{2}}\mathbf{u} - \mathbf{b}\cdot A^{-\frac{1}{2}}\mathbf{u} - cp,\, \nabla\cdot A^{\frac{1}{2}}\mathbf{v} - \mathbf{b}\cdot A^{-\frac{1}{2}}\mathbf{v} - c\xi\rangle \\ & + & \langle \nabla\times A^{-\frac{1}{2}}\mathbf{u},\, \nabla\times A^{-\frac{1}{2}}\mathbf{v}\rangle \end{array} \tag{4.7}$$

and

$$\begin{aligned} &a_2^*(\mathbf{w}, r, s;\, \mathbf{v}, \xi, \eta) := \langle \mathcal{L}_2^*(\mathbf{w}, r, s)^t,\, \mathcal{L}_2^*(\mathbf{v}, \xi, \eta)^t\rangle \\ &\qquad = \langle \mathbf{w} - A^{\frac{1}{2}}\nabla r - A^{-\frac{1}{2}}\mathbf{b} r + A^{-\frac{1}{2}}\nabla^{\perp} s,\, \mathbf{v} - A^{\frac{1}{2}}\nabla\xi - A^{-\frac{1}{2}}\mathbf{b}\xi + A^{-\frac{1}{2}}\nabla^{\perp}\eta\rangle \\ &\qquad + \langle \nabla\cdot A^{\frac{1}{2}}\mathbf{w} - cr,\, \nabla\cdot A^{\frac{1}{2}}\mathbf{v} - c\xi\rangle + \langle \nabla\times A^{-\frac{1}{2}}\mathbf{w},\, \nabla\times A^{-\frac{1}{2}}\mathbf{v}\rangle, \end{aligned} \tag{4.8}$$

and the linear functional

$$f_2(\mathbf{v}, \xi, \eta) := \langle (\mathbf{0}, -f, 0)^t,\, (\mathbf{v}, \xi, \eta)^t\rangle = -\langle f, \xi\rangle. \tag{4.9}$$

In this context, the variational problem described in (2.10) becomes that of finding $(\mathbf{w}, r, s) \in \mathcal{D}$ such that

$$a_2^*(\mathbf{w}, r, s;\, \mathbf{v}, \xi, \eta) = f_2(\mathbf{v}, \xi, \eta) \quad \forall\ (\mathbf{v}, \xi, \eta) \in \mathcal{D}. \tag{4.10}$$

This problem is the core of what, for later reference, we call the FOSLL$_e^*$ method for solving (1.5) in two dimensions.

In what follows, C denotes a generic constant that depends only on A, $\mathbf{b}$, c, and Ω, but it may change meaning with every occurrence.

THEOREM 4.1. *Operators $\mathcal{L}_2$ and $\mathcal{L}_2^*$ are bijective, and bilinear forms $a_2(\,\cdot\,;\,\cdot\,)$ and $a_2^*(\,\cdot\,;\,\cdot\,)$ are coercive and continuous on $\mathcal{D} = \mathcal{D}^*$; that is, there exist positive constants α_0, α_1 and α_0^*, α_1^*, which depend only on A, $\mathbf{b}$, c, and Ω, such that*

$$\alpha_0 \,\|(\mathbf{u}, p, q)\|_{\mathcal{D}}^2 \leq a_2(\mathbf{u}, p, q;\, \mathbf{u}, p, q) \leq \alpha_1 \,\|(\mathbf{u}, p, q)\|_{\mathcal{D}}^2\,, \tag{4.11}$$

$$\alpha_0^* \,\|(\mathbf{u}, p, q)\|_{\mathcal{D}}^2 \leq a_2^*(\mathbf{u}, p, q;\, \mathbf{u}, p, q) \leq \alpha_1^* \,\|(\mathbf{u}, p, q)\|_{\mathcal{D}}^2 \tag{4.12}$$

for any $(\mathbf{u}, p, q) \in \mathcal{D}$. Furthermore, problem (4.10) has a unique solution $(\mathbf{w}, r, s) \in \mathcal{D}$ and it satisfies the a priori estimate

$$\|(\mathbf{w}, r, s)\|_{\mathcal{D}} \leq C \|f\|_{-1,D}. \tag{4.13}$$

Proof. Continuity of the bilinear forms follows from repeated use of the triangle inequality. To establish coercivity of $a_2(\,\cdot\,;\,\cdot\,)$, we use the result from [11] that $a_2(\mathbf{u}, p, 0;\, \mathbf{u}, p, 0)$ is coercive for $(\mathbf{u}, p) \in \mathcal{W}_{ND}(A) \times H_D^1$. This result and the triangle inequality yield

$$\begin{aligned}
\|\mathbf{u}\|_{\mathcal{W}(A)}^2 + \|p\|_1^2 &\leq C \left(\|\mathbf{u} - A^{\frac{1}{2}} \nabla p\|^2 + \|\nabla \cdot A^{\frac{1}{2}} \mathbf{u} - \mathbf{b} \cdot A^{-\frac{1}{2}} \mathbf{u} - cp\|^2 + \|\nabla \times A^{-\frac{1}{2}} \mathbf{u}\|^2 \right) \\
&\leq C \Big(\|\mathbf{u} - A^{\frac{1}{2}} \nabla p + A^{-\frac{1}{2}} \nabla^{\perp} q\|^2 + \|A^{-\frac{1}{2}} \nabla^{\perp} q\|_1^2 \\
&\qquad + \|\nabla \cdot A^{\frac{1}{2}} \mathbf{u} - \mathbf{b} \cdot A^{-\frac{1}{2}} \mathbf{u} - cp\|^2 + \|\nabla \times A^{-\frac{1}{2}} \mathbf{u}\|^2 \Big) \\
&\leq C \left(a_2(\mathbf{u}, p, q;\, \mathbf{u}, p, q) + |q|_1^2 \right)
\end{aligned}$$

for $(\mathbf{u}, p, q)^t \in \mathcal{D}$. Now, for $q \in H_N^1$, we have

$$\begin{aligned}
\langle A^{-\frac{1}{2}} \nabla^{\perp} q, A^{-\frac{1}{2}} \nabla^{\perp} q \rangle &= \langle A^{-\frac{1}{2}} \nabla^{\perp} q,\, \mathbf{u} - A^{\frac{1}{2}} \nabla p + A^{-\frac{1}{2}} \nabla^{\perp} q \rangle - \langle A^{-\frac{1}{2}} \nabla^{\perp} q,\, \mathbf{u} \rangle \\
&= \langle A^{-\frac{1}{2}} \nabla^{\perp} q,\, \mathbf{u} - A^{\frac{1}{2}} \nabla p + A^{-\frac{1}{2}} \nabla^{\perp} q \rangle - \langle q,\, \nabla \times A^{-\frac{1}{2}} \mathbf{u} \rangle \\
&\leq \|A^{-\frac{1}{2}} \nabla^{\perp} q\| \|\mathbf{u} - A^{\frac{1}{2}} \nabla p + A^{-\frac{1}{2}} \nabla^{\perp} q\| + \|q\| \|\nabla \times A^{-\frac{1}{2}} \mathbf{u}\|,
\end{aligned}$$

which, together with the bound $\|q\| \leq C|q|_1 \leq C\|A^{-\frac{1}{2}} \nabla^{\perp} q\|$, yields

$$\|q\|_1 \leq C \|A^{-\frac{1}{2}} \nabla^{\perp} q\| \leq C a_2(\mathbf{u}, p, q;\, \mathbf{u}, p, q)^{\frac{1}{2}}.$$

Combining the above results yields coercivity of $a_2(\,\cdot\,;\,\cdot\,)$.

We next show that $\mathcal{L}_2^*$ is injective. Assume that $\mathcal{L}_2^*(\mathbf{w}, r, s)^t = (\mathbf{0}, 0, 0)^t$ for some $(\mathbf{w}, r, s)^t \in \mathcal{D}$. This implies that

$$\begin{aligned}
\mathbf{w} - A^{\frac{1}{2}} \nabla r - A^{-\frac{1}{2}} \mathbf{b} r + A^{-\frac{1}{2}} \nabla^{\perp} s &= \mathbf{0}, \\
\nabla \cdot A^{\frac{1}{2}} \mathbf{w} - cr &= 0, \\
\nabla \times A^{-\frac{1}{2}} \mathbf{w} &= 0.
\end{aligned}$$

Multiplying the first equation by $\mathbf{n} \cdot A^{\frac{1}{2}}$ and using boundary conditions (4.6) confirm that r satisfies the adjoint boundary conditions in (1.6). Solving for $\mathbf{w}$ in the first equation and substituting into the second equation confirm that r satisfies homogenous adjoint problem (1.6). By assumption, $r = 0$. Now, multiplying the first equation by $\boldsymbol{\tau} \cdot A^{-\frac{1}{2}}$ and using (4.6) again yield

$$\boldsymbol{\tau} \cdot A^{-1} \nabla^{\perp} s = \frac{1}{\det(A)} \mathbf{n} \cdot A \nabla s = 0$$

on Γ_D. Combining the first and third equations, together with (4.6), yields

$$\nabla \times A^{-1} \nabla^{\perp} s = \nabla \cdot \frac{1}{\det(A)} A \nabla s = 0.$$

Together with the boundary condition $s = 0$ on Γ_N, this yields $s = 0$, which in turn implies $\mathbf{w} = \mathbf{0}$. Thus, $\mathcal{L}_2^*$ is injective.

Now, coercivity of $a_2(\,\cdot\,;\,\cdot\,)$ implies that $\mathcal{L}_2^{-1}$ is continuous and $\mathcal{R}$ is closed. Since $\mathcal{L}_2^*$ is injective, that is, it has no null space, then $\mathcal{L}_2$ is bijective: $\mathcal{R} = \mathcal{V}_2$. Application of Lemma 2.1 yields coercivity of $a_2^*(\,\cdot\,;\,\cdot\,)$. We note that this also implies that $\mathcal{L}_2^*$ is bijective.

It is easy to see that linear form $f_2(\cdot)$ is continuous on $\mathcal{D}$. Hence, by the Lax–Milgram lemma [15], problem (4.10) has a unique solution $(\mathbf{w}, r, s) \in \mathcal{D}$. Derivation of the a priori estimate is straightforward. ∎

The result that $\mathcal{L}_2^*$ is bijective implies that the solution of $\mathcal{L}_2(\mathbf{u}, p, q)^t = (\mathbf{0}, f, 0)^t$ is in the range of $\mathcal{L}_2^*$.

THEOREM 4.2. *Let* $(\mathbf{w}, r, s) \in \mathcal{D}$ *be the solution of problem* (4.10) *and let*

$$\mathbf{u} = \mathbf{w} - A^{\frac{1}{2}} \nabla r - A^{-\frac{1}{2}} \mathbf{b} r + A^{-\frac{1}{2}} \nabla^{\perp} s, \quad p = \nabla \cdot A^{\frac{1}{2}} \mathbf{w} - cr, \quad q = \nabla \times A^{-\frac{1}{2}} \mathbf{w}.$$

Then p is the solution of problem (1.5), $A^{\frac{1}{2}} \mathbf{u}$ *is the flux, and* $q = 0$.

Proof. The proof follows immediately from Theorem 4.1. Let p satisfy (1.5) and let $\mathbf{u} = A^{\frac{1}{2}} \nabla p$ and $q = 0$. Then $\mathcal{L}_2(\mathbf{u}, p, q)^t = (\mathbf{0}, f, 0)^t$. Since both $\mathcal{L}_2$ and $\mathcal{L}_2^*$ are bijective, then $(\mathbf{u}, p, q)^t$ is unique and in the range of $\mathcal{L}_2^*$.

Coercivity of $\mathcal{L}_2^*$ implies that the solution of (4.10), say $(\mathbf{w}, r, s)^t \in \mathcal{D}$, is the unique minimizer of $G_e(\mathbf{v}, \xi, \eta) := \|\mathcal{L}_2^*(\mathbf{v}, \xi, \eta)^t - (\mathbf{u}, p, q)^t\|^2$ over $\mathcal{D}$, which implies that $\mathcal{L}_2^*(\mathbf{w}, r, s)^t = (\mathbf{u}, p, q)^t$. ∎

Remark 4.1. When $\mathbf{b} = \mathbf{0}$, the solution, $(\mathbf{w}, r, s)^t$, of (4.10) will have $s = 0$. To see this, note that the system $\mathcal{L}_2^*(\mathbf{w}, r, s)^t = (\mathbf{u}, p, q)^t$ in this case becomes

$$\begin{aligned} \mathbf{w} - A^{\frac{1}{2}} \nabla r + A^{-\frac{1}{2}} \nabla^{\perp} s &= \mathbf{u} = A^{\frac{1}{2}} \nabla p, \\ \nabla \cdot A^{\frac{1}{2}} \mathbf{w} - cr &= p, \\ \nabla \times A^{-\frac{1}{2}} \mathbf{w} &= 0. \end{aligned}$$

As in the proof of Theorem 4.1, multiplying the first equation by $\boldsymbol{\tau} \cdot A^{-\frac{1}{2}}$ and using boundary conditions (4.6) yield $\mathbf{n} \cdot A \nabla s = 0$ on Γ_D. Substituting the first equation into the last equation yields $\nabla \times A^{-1} \nabla^{\perp} s = \nabla \cdot \frac{1}{\det(A)} A \nabla s = 0$. But $s = 0$ on Γ_N, so this implies that $s = 0$.

To gain insight into the effectiveness of this FOSLL$_e^*$ approach, we consider the square system

$$\mathcal{L}_2 \mathcal{L}_2^* (\mathbf{w}, r, s)^t = (\mathbf{0}, -f, 0)^t \quad \text{in } \Omega\,, \tag{4.14}$$

where

$$\mathcal{L}_2\mathcal{L}_2^* = ((\mathcal{L}_2\mathcal{L}_2^*)_1, (\mathcal{L}_2\mathcal{L}_2^*)_2, (\mathcal{L}_2\mathcal{L}_2^*)_3) \;, \tag{4.15}$$

with

$$(\mathcal{L}_2\mathcal{L}_2^*)_1 = \begin{pmatrix} I - A^{\frac{1}{2}}\nabla\nabla\cdot A^{\frac{1}{2}} + A^{-\frac{1}{2}}\nabla^{\perp}\nabla\times A^{-\frac{1}{2}} \\ \nabla\cdot A^{\frac{1}{2}} - \mathbf{b}\cdot A^{-\frac{1}{2}} - c\nabla\cdot A^{\frac{1}{2}} \\ \nabla\times A^{-\frac{1}{2}} \end{pmatrix},$$

$$(\mathcal{L}_2\mathcal{L}_2^*)_2 = \begin{pmatrix} -A^{\frac{1}{2}}\nabla - A^{-\frac{1}{2}}\mathbf{b} + A^{\frac{1}{2}}\nabla c \\ -\nabla\cdot A\nabla + \left(\mathbf{b}\cdot A^{-1}\mathbf{b} - (\nabla\cdot\mathbf{b}) + c^2\right) I \\ -\nabla\times A^{-1}\mathbf{b} \end{pmatrix},$$

$$(\mathcal{L}_2\mathcal{L}_2^*)_3 = \begin{pmatrix} A^{-\frac{1}{2}}\nabla^{\perp} \\ -\mathbf{b}\cdot A^{-1}\nabla^{\perp} \\ \nabla\times A^{-1}\nabla^{\perp} \end{pmatrix}.$$

By $(\nabla\cdot\mathbf{b})I$, we mean the 0th-degree operator defined by $(\nabla\cdot\mathbf{b})Ir = r\nabla\cdot\mathbf{b}$. Note that this term arises because $\mathbf{b}\cdot\nabla r - \nabla\cdot(r\mathbf{b}) = r\nabla\cdot\mathbf{b}$. The diagonal of $\mathcal{L}_2\mathcal{L}_2^*$ consists of second-order operators, each of which is locally H^1-elliptic, while the off-diagonal terms are all first-order operators. We say that the formal normal is *differentially diagonally dominant*. The size of the off-diagonal terms grows with the size of $\mathbf{b}$, but the effect is dependent on the product $|\mathbf{b}|h$ for some mesh scale h. This means that the coupling between variables is negligible at fine-grid scales and that the conflict is only between smooth error components on coarser grids. This, in turn, implies that W-cycles can achieve efficiency that is fairly insensitive to the size of $|\mathbf{b}|$, as the results in section 7 show.

4.2. Three dimensions. In three dimensions, $\nabla\times$ is no longer a scalar and its formal adjoint is not $\nabla^{\perp}$ but $\nabla\times$. Moreover, $\nabla\times A^{-1}\nabla\times$ is a singular operator. We need to add a vector slack variable, $\mathbf{v}$, in addition to the scalar slack variable, q. Consider the following extended first-order system:

$$\left\{\begin{array}{rcll} \mathbf{u} - A^{\frac{1}{2}}\nabla p + A^{-\frac{1}{2}}\nabla\times\mathbf{v} &=& \mathbf{0} & \text{in} \quad \Omega, \\ -\nabla\cdot A^{\frac{1}{2}}\mathbf{u} + \mathbf{b}\cdot A^{-\frac{1}{2}}\mathbf{u} + cp &=& f & \text{in} \quad \Omega, \\ \nabla\times A^{-\frac{1}{2}}\mathbf{u} - \mathbf{v} - \nabla q &=& \mathbf{0} & \text{in} \quad \Omega, \\ \nabla\cdot\mathbf{v} &=& 0 & \text{in} \quad \Omega \end{array}\right. \tag{4.16}$$

with boundary conditions

$$\left\{\begin{array}{rcll} \mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{u} &=& 0 & \text{on} \quad \Gamma_N, \\ \mathbf{n}\times A^{-\frac{1}{2}}\mathbf{u} &=& \mathbf{0} & \text{on} \quad \Gamma_D, \\ p &=& 0 & \text{on} \quad \Gamma_D, \\ \mathbf{n}\cdot\mathbf{v} &=& 0 & \text{on} \quad \Gamma_D, \\ \mathbf{n}\times\mathbf{v} &=& \mathbf{0} & \text{on} \quad \Gamma_N, \\ q &=& 0 & \text{on} \quad \Gamma_N. \end{array}\right. \tag{4.17}$$

Again changing the sign of the second equation for convenience, the differential operator for system (4.16) is

$$\mathcal{L}_3 = \begin{pmatrix} I & -A^{\frac{1}{2}}\nabla & A^{-\frac{1}{2}}\nabla\times & 0 \\ \nabla\cdot A^{\frac{1}{2}} - \mathbf{b}\cdot A^{-\frac{1}{2}} & -cI & 0 & 0 \\ \nabla\times A^{-\frac{1}{2}} & 0 & -I & -\nabla \\ 0 & 0 & \nabla\cdot & 0 \end{pmatrix}$$

and its formal adjoint is

$$\mathcal{L}_3^* = \begin{pmatrix} I & -A^{\frac{1}{2}}\nabla - A^{-\frac{1}{2}}\mathbf{b} & A^{-\frac{1}{2}}\nabla\times & 0 \\ \nabla\cdot A^{\frac{1}{2}} & -cI & 0 & 0 \\ \nabla\times A^{-\frac{1}{2}} & 0 & -I & -\nabla \\ 0 & 0 & \nabla\cdot & 0 \end{pmatrix}.$$

In this context, $\mathcal{L}_3 : \mathcal{D} \subseteq \mathcal{V}_1 \to \mathcal{V}_2$, where $\mathcal{V}_1 = \mathcal{V}_2 = \mathcal{V} := L^2(\Omega)^8$, and $\mathcal{D} = \mathcal{W}_{ND}(A) \times H_D^1 \times \mathcal{W}_{DN}(I) \times H_N^1$, which is a Hilbert space under the product norm

$$\|(\mathbf{u}, p, \mathbf{v}, q)\|_{\mathcal{D}} := \Big(\|\mathbf{u}\|_{\mathcal{W}(A)}^2 + \|p\|_1^2 + \|\mathbf{v}\|_{\mathcal{W}}^2 + \|q\|_1^2\Big)^{\frac{1}{2}} \tag{4.18}$$

and is compactly embedded in $\mathcal{V}$.

The domain of $\mathcal{L}_3^*$ involves the *dual variables* $(\mathbf{w}, r, \mathbf{x}, s)^t$ with $\mathbf{w} \in H(\operatorname{div} A^{\frac{1}{2}}; \Omega) \cap H(\mathbf{curl}\, A^{-\frac{1}{2}}; \Omega)$, $\mathbf{x} \in H(\operatorname{div}; \Omega) \cap H(\mathbf{curl}; \Omega)$, and r, $s \in H^1(\Omega)$. To determine the boundary conditions associated with $\mathcal{L}_3^*$, we compute

$$\begin{aligned}
&\langle \mathcal{L}_3\,(\mathbf{u},\, p,\, \mathbf{v},\, q)^t,\, (\mathbf{w},\, r,\, \mathbf{x},\, s)^t \rangle \\
&\quad = \langle (\mathbf{u},\, p,\, \mathbf{v},\, q)^t,\, \mathcal{L}_3^*\,(\mathbf{w},\, r,\, \mathbf{x},\, s)^t \rangle - \textstyle\int_{\partial\Omega} p(\mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{w}) + \int_{\partial\Omega} \mathbf{v}\cdot(\mathbf{n}\times A^{-\frac{1}{2}}\mathbf{w}) \\
&\qquad + \textstyle\int_{\partial\Omega} (\mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{u}) r - \int_{\partial\Omega} (\mathbf{n}\times A^{-\frac{1}{2}}\mathbf{u})\cdot\mathbf{x} - \int_{\partial\Omega} q(\mathbf{n}\cdot\mathbf{x}) + \int_{\partial\Omega} (\mathbf{n}\cdot\mathbf{v}) s \\
&\quad = \langle (\mathbf{u},\, p,\, \mathbf{v},\, q)^t,\, \mathcal{L}_3^*\,(\mathbf{w},\, r,\, \mathbf{x},\, s)^t \rangle - \textstyle\int_{\Gamma_N} p(\mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{w}) + \int_{\Gamma_D} \mathbf{v}\cdot(\mathbf{n}\times A^{-\frac{1}{2}}\mathbf{w}) \\
&\qquad + \textstyle\int_{\Gamma_D} (\mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{u}) r - \int_{\Gamma_N} (\mathbf{n}\times A^{-\frac{1}{2}}\mathbf{u})\cdot\mathbf{x} - \int_{\Gamma_D} q(\mathbf{n}\cdot\mathbf{x}) + \int_{\Gamma_N} (\mathbf{n}\cdot\mathbf{v}) s,
\end{aligned}$$

where (4.17) was used to reduce the boundary integral terms. The remaining boundary integrals vanish for every $(\mathbf{u}, p, \mathbf{v}, q)^t \in \mathcal{D}$ if and only if we enforce the boundary conditions

$$\left\{ \begin{array}{rcll} \mathbf{n}\cdot A^{\frac{1}{2}}\mathbf{w} & = & 0 & \text{on} \quad \Gamma_N, \\ \mathbf{n}\times A^{-\frac{1}{2}}\mathbf{w} & = & \mathbf{0} & \text{on} \quad \Gamma_D, \\ r & = & 0 & \text{on} \quad \Gamma_D, \\ \mathbf{n}\cdot\mathbf{x} & = & 0 & \text{on} \quad \Gamma_D, \\ \mathbf{n}\times\mathbf{x} & = & \mathbf{0} & \text{on} \quad \Gamma_N, \\ s & = & 0 & \text{on} \quad \Gamma_N. \end{array} \right. \tag{4.19}$$

Comparing (4.19) with (4.17), we see that $\mathcal{D}^* = \mathcal{D}$.

We next define the bilinear forms

(4.20)

$$\begin{aligned}
a_3(\mathbf{u}, p, \mathbf{v}, q;\, \mathbf{y}, \xi, \mathbf{z}, \eta) &:= \langle \mathcal{L}_3(\mathbf{u}, p, \mathbf{v}, q)^t,\, \mathcal{L}_3(\mathbf{y}, \xi, \mathbf{z}, \eta)^t \rangle \\
&= \langle \mathbf{u} - A^{\frac{1}{2}}\nabla p + A^{-\frac{1}{2}}\nabla\times\mathbf{v},\, \mathbf{y} - A^{\frac{1}{2}}\nabla\xi + A^{-\frac{1}{2}}\nabla\times\mathbf{z} \rangle \\
&+ \langle \nabla\cdot A^{\frac{1}{2}}\mathbf{u} - \mathbf{b}\cdot A^{-\frac{1}{2}}\mathbf{u} - cp,\, \nabla\cdot A^{\frac{1}{2}}\mathbf{y} - \mathbf{b}\cdot A^{-\frac{1}{2}}\mathbf{y} - c\xi \rangle \\
&+ \langle \nabla\times A^{-\frac{1}{2}}\mathbf{u} - \mathbf{v} - \nabla q,\, \nabla\times A^{-\frac{1}{2}}\mathbf{y} - \mathbf{z} - \nabla\eta \rangle + \langle \nabla\cdot\mathbf{v},\, \nabla\cdot\mathbf{z} \rangle
\end{aligned}$$

and

$$\begin{aligned}
&a_3^*(\mathbf{w}, r, \mathbf{x}, s;\, \mathbf{y}, \xi, \mathbf{z}, \eta) := \langle \mathcal{L}_3^*(\mathbf{w}, r, \mathbf{x}, s)^t,\, \mathcal{L}_3^*(\mathbf{y}, \xi, \mathbf{z}, \eta)^t \rangle \\
&\quad = \langle \mathbf{w} - A^{\frac{1}{2}}\nabla r - A^{-\frac{1}{2}}\mathbf{b}r + A^{-\frac{1}{2}}\nabla\times\mathbf{x},\, \mathbf{y} - A^{\frac{1}{2}}\nabla\xi - A^{-\frac{1}{2}}\mathbf{b}\xi + A^{-\frac{1}{2}}\nabla\times\mathbf{z} \rangle \\
&\quad + \langle \nabla\cdot A^{\frac{1}{2}}\mathbf{w} - cr,\, \nabla\cdot A^{\frac{1}{2}}\mathbf{y} - c\xi \rangle \\
&\quad + \langle \nabla\times A^{-\frac{1}{2}}\mathbf{w} - \mathbf{x} - \nabla s,\, \nabla\times A^{-\frac{1}{2}}\mathbf{y} - \mathbf{z} - \nabla\eta \rangle + \langle \nabla\cdot\mathbf{x},\, \nabla\cdot\mathbf{z} \rangle
\end{aligned} \tag{4.21}$$

and the linear functional

$$f_3(\mathbf{y}, \xi, \mathbf{z}, \eta) := \langle (\mathbf{0}, -f, \mathbf{0}, 0)^t,\ (\mathbf{y}, \xi, \mathbf{z}, \eta)^t \rangle = -\langle f, \xi \rangle. \tag{4.22}$$

In this context, the variational problem described in (2.10) becomes that of finding $(\mathbf{w}, r, \mathbf{x}, s) \in \mathcal{D}$ such that

$$a_3^*(\mathbf{w}, r, \mathbf{x}, s;\ \mathbf{y}, \xi, \mathbf{z}, \eta) = f_3(\mathbf{y}, \xi, \mathbf{z}, \eta) \quad \forall\ (\mathbf{y}, \xi, \mathbf{z}, \eta) \in \mathcal{D}. \tag{4.23}$$

This problem is the core of what, for later reference, we call the FOSLL$_e^*$ method for solving (1.5) in three dimensions.

THEOREM 4.3. *Operators $\mathcal{L}_3$ and $\mathcal{L}_3^*$ are bijective, and bilinear forms $a_3(\,\cdot\,;\,\cdot\,)$ and $a_3^*(\,\cdot\,;\,\cdot\,)$ are coercive and continuous on $\mathcal{D} = \mathcal{D}^*$; that is, there exist positive constants α_0, α_1 and α_0^*, α_1^*, which depend only on A, $\mathbf{b}$, c, and Ω, such that*

$$\alpha_0 \|(\mathbf{u}, p, \mathbf{v}, q)\|_{\mathcal{D}}^2 \le a_3(\mathbf{u}, p, \mathbf{v}, q;\ \mathbf{u}, p, \mathbf{v}, q) \le \alpha_1 \|(\mathbf{u}, p, \mathbf{v}, q)\|_{\mathcal{D}}^2\,, \tag{4.24}$$

$$\alpha_0^* \|(\mathbf{u}, p, \mathbf{v}, q)\|_{\mathcal{D}}^2 \le a_3^*(\mathbf{u}, p, \mathbf{v}, q;\ \mathbf{u}, p, \mathbf{v}, q) \le \alpha_1^* \|(\mathbf{u}, p, \mathbf{v}, q)\|_{\mathcal{D}}^2 \tag{4.25}$$

for any $(\mathbf{u}, p, \mathbf{v}, q) \in \mathcal{D}$. Furthermore, problem (4.23) has a unique solution $(\mathbf{w}, r, \mathbf{x}, s) \in \mathcal{D}$ and it satisfies the a priori estimate

$$\|(\mathbf{w}, r, \mathbf{x}, s)\|_{\mathcal{D}} \le C \|f\|_{-1,D}. \tag{4.26}$$

Proof. Continuity of the forms follows from repeated use of the triangle inequality. The remainder of the proof follows that of Theorem 4.1. To establish coercivity of $a_3(\,\cdot\,;\,\cdot\,)$, first note that, for any $(\mathbf{u}, p, \mathbf{v}, q) \in \mathcal{D}$, integration by parts gives

$$\begin{aligned}\langle \mathbf{u}, A^{-\frac{1}{2}} \nabla\times\mathbf{v} \rangle &= \langle \nabla\times A^{-\frac{1}{2}}\mathbf{u}, \mathbf{v} \rangle + \textstyle\int_{\Gamma_D} \left(\mathbf{n} \times A^{-\frac{1}{2}}\mathbf{u}\right) \cdot \mathbf{v} - \int_{\Gamma_N} (A^{-\frac{1}{2}}\mathbf{u}) \cdot (\mathbf{n} \times \mathbf{v}) \\ &= \langle \nabla\times A^{-\frac{1}{2}}\mathbf{u}, \mathbf{v} \rangle,\end{aligned}$$

$$\langle A^{\frac{1}{2}} \nabla p, A^{-\frac{1}{2}} \nabla\times\mathbf{v} \rangle = \textstyle\int_{\Gamma_D} \mathbf{v} \cdot (\mathbf{n} \times \nabla p) - \int_{\Gamma_N} (\mathbf{n} \times \mathbf{v}) \cdot \nabla p = 0,$$

$$\langle \nabla\times A^{-\frac{1}{2}}\mathbf{u}, \nabla q \rangle = -\textstyle\int_{\Gamma_D} \left(\mathbf{n} \times A^{-\frac{1}{2}}\mathbf{u}\right) \cdot \nabla q + \int_{\Gamma_N} \left(A^{-\frac{1}{2}}\mathbf{u}\right) \cdot (\mathbf{n} \times \nabla q) = 0.$$

Next, for $q \in H_N^1$, we have

$$\begin{aligned}\langle \nabla q,\ \nabla q \rangle &= \langle \nabla q,\ \nabla q + \mathbf{v} - \nabla\times A^{-\frac{1}{2}}\mathbf{u} \rangle - \langle \nabla q, \mathbf{v} \rangle \\ &\le \|\nabla q\| \|\nabla\times A^{-\frac{1}{2}}\mathbf{u} - \mathbf{v} - \nabla q\| + \|q\| \|\nabla \cdot \mathbf{v}\|,\end{aligned}$$

which, together with the bound $\|q\| \le C\|\nabla q\|$, yields

$$\|q\| \le C\|\nabla q\| \le C a_3(\mathbf{u}, p, \mathbf{v}, q\,;\, \mathbf{u}, p, \mathbf{v}, q)^{\frac{1}{2}}. \tag{4.27}$$

A simple calculation yields

$$\begin{aligned}a_3(\mathbf{u}, p, \mathbf{v}, q\,;\, \mathbf{u}, p, \mathbf{v}, q) &= \|\mathbf{u} - A^{\frac{1}{2}} \nabla p + A^{-\frac{1}{2}} \nabla\times\mathbf{v}\|^2 \\ &\quad + \|\nabla \cdot A^{\frac{1}{2}}\mathbf{u} - \mathbf{b} \cdot A^{-\frac{1}{2}}\mathbf{u} - cp\|^2 + \|\nabla\times A^{-\frac{1}{2}}\mathbf{u} - \mathbf{v} - \nabla q\|^2 + \|\nabla \cdot \mathbf{v}\|^2 \\ &= \|\mathbf{u} - A^{\frac{1}{2}} \nabla p\|^2 + \|A^{-\frac{1}{2}} \nabla\times\mathbf{v}\|^2 + \|\nabla \cdot A^{\frac{1}{2}}\mathbf{u} - \mathbf{b} \cdot A^{-\frac{1}{2}}\mathbf{u} - cp\|^2 \\ &\quad + \|\nabla\times A^{-\frac{1}{2}}\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2(\nabla \cdot \mathbf{v}, q) + \|\nabla q\|^2 + \|\nabla \cdot \mathbf{v}\|^2\end{aligned} \tag{4.28}$$

for $(\mathbf{u}, p, \mathbf{v}, q) \in \mathcal{D}$. We again use the result from [11] that $a_3(\mathbf{u}, p, \mathbf{0}, 0;\, \mathbf{u}, p, \mathbf{0}, 0)$ is coercive for $(\mathbf{u}, p) \in \mathcal{W}_{ND}(A) \times H^1_D$. This yields

(4.29)
$$|\mathbf{u}\|^2_{\mathcal{W}(A)} + \|p\|^2_1 \le C\left(\|\mathbf{u} - A^{\frac{1}{2}}\nabla p\|^2 + \|\nabla\cdot A^{\frac{1}{2}}\mathbf{u} - \mathbf{b}\cdot A^{-\frac{1}{2}}\mathbf{u} - cp\|^2 + \|\nabla\times A^{-\frac{1}{2}}\mathbf{u}\|^2\right)$$

for $(\mathbf{u}, p, \mathbf{0}, 0)^t \in \mathcal{D}$. Combining (4.27), (4.28), and (4.30) yields

$$\begin{aligned}\|\mathbf{u}\|^2_{\mathcal{W}(A)} + \|p\|^2_1 + \|\mathbf{v}\|^2_{\mathcal{W}} + \|q\|^2_1 &\le C\left(a_3(\mathbf{u}, p, \mathbf{v}, q\,;\, \mathbf{u}, p, \mathbf{v}, q) + \|q\|\|\nabla\cdot\mathbf{v}\|\right)\\ &\le C a_3(\mathbf{u}, p, \mathbf{v}, q\,;\, \mathbf{u}, p, \mathbf{v}, q)\end{aligned}$$

for $(\mathbf{u}, p, \mathbf{v}, q)^t \in \mathcal{D}$, which confirms coercivity of $a_3(\,\cdot\,;\,\cdot\,)$.

We next show that $\mathcal{L}_3^*$ is injective. Assume that $\mathcal{L}_3^*(\mathbf{w}, r, \mathbf{x}, s)^t = (\mathbf{0}, 0, \mathbf{0}, 0)^t$ for some $(\mathbf{w}, r, \mathbf{x}, s)^t \in \mathcal{D}$. This implies that

$$\begin{array}{rcl}\mathbf{w} - A^{\frac{1}{2}}\nabla r - A^{-\frac{1}{2}}\mathbf{b}r + A^{-\frac{1}{2}}\nabla\times\mathbf{x} &=& \mathbf{0},\\ \nabla\cdot A^{\frac{1}{2}}\mathbf{w} - cr &=& 0,\\ \nabla\times A^{-\frac{1}{2}}\mathbf{w} - \mathbf{x} - \nabla s &=& 0,\\ \nabla\cdot\mathbf{x} &=& 0.\end{array} \tag{4.30}$$

First, note that $\mathbf{n}\times\mathbf{x} = \mathbf{0}$ on Γ_N implies that $\mathbf{n}\cdot\nabla\times\mathbf{x} = 0$ on Γ_N. Multiplying the first equation by $\mathbf{n}\cdot A^{\frac{1}{2}}$ and using boundary conditions (4.19) confirm that r satisfies the adjoint boundary conditions in (1.6). Eliminating $\mathbf{w}$ from the first two equations confirms that r satisfies homogenous adjoint problem (1.6). By assumption, $r = 0$.

Again, notice that $\mathbf{n}\times A^{-\frac{1}{2}}\mathbf{w} = 0$ on Γ_D implies that $\mathbf{n}\cdot\nabla\times A^{-\frac{1}{2}}\mathbf{w} = 0$ on Γ_D. Multiplying the third equation by $\mathbf{n}\cdot$ shows that $\mathbf{n}\cdot\nabla s = 0$ on Γ_D. Eliminating $\mathbf{x}$ from the last two equations yields $\nabla\cdot\nabla s = 0$. Together with the condition $s = 0$ on Γ_N from (4.19), we conclude that $s = 0$.

Finally, we use the first and third equations with $r = s = 0$ to get

$$\langle\mathbf{w},\, \mathbf{w}\rangle = -\langle\mathbf{w},\, A^{-\frac{1}{2}}\nabla\times\mathbf{x}\rangle = -\langle\nabla\times A^{-\frac{1}{2}}\mathbf{w},\, \mathbf{x}\rangle = -\langle\mathbf{x},\, \mathbf{x}\rangle,$$

which implies that $\mathbf{w} = \mathbf{x} = \mathbf{0}$. Thus, $\mathcal{L}_3^*$ is injective.

Now, coercivity of $a_3(\,\cdot\,;\,\cdot\,)$ implies that $\mathcal{L}_3^{-1}$ is continuous and that $\mathcal{R}$ is closed. Since $\mathcal{L}_3^*$ is injective, that is, it has no null space, then $\mathcal{L}_3$ is bijective: $\mathcal{R} = \mathcal{V}$. Application of Lemma 2.1 yields coercivity of $\mathcal{L}_3^*$. We note that this also implies that $\mathcal{L}_3^*$ is bijective.

It is easy to see that linear form $f_3(\cdot)$ is continuous in $\mathcal{D}$. Hence, by the Lax–Milgram lemma [15], problem (4.23) has a unique solution $(\mathbf{w}, r, \mathbf{x}, s) \in \mathcal{D}$. Derivation of the a priori estimate is straightforward. ◻

Remark 4.2. If both Γ_D and Γ_N are connected, then the inequality

$$\|\mathbf{v}\|^2 \le C\left(\|\nabla\cdot\mathbf{v}\|^2 + \|\nabla\times\mathbf{v}\|^2\right) \tag{4.31}$$

holds for every $\mathbf{v} \in \mathcal{W}_{DN}$. In this case, the result of Theorem 4.3 still holds when we remove $-\mathbf{v}$ from the third equation of (4.16). This has the effect of removing $-I$ from the $(3,3)$ position of $\mathcal{L}_3$ and $\mathcal{L}_3^*$.

The result that $\mathcal{L}_3^*$ is bijective implies that the solution of $\mathcal{L}_3(\mathbf{u}, p, \mathbf{v}, q)^t = (\mathbf{0}, f, \mathbf{0}, 0)^t$ is in the range of $\mathcal{L}_3^*$. This yields the following result.

THEOREM 4.4. *Let* $(\mathbf{w}, r, \mathbf{x}, s) \in \mathcal{D}$ *be the solution of problem* (4.23) *and let*

$$\begin{aligned} \mathbf{u} &= \mathbf{w} - A^{\frac{1}{2}}\nabla r - A^{-\frac{1}{2}}\mathbf{b}r + A^{-\frac{1}{2}}\nabla\times\mathbf{x}, & p &= \nabla\cdot A^{\frac{1}{2}}\mathbf{w} - cr,\\ \mathbf{v} &= \nabla\times(A^{-\frac{1}{2}}\mathbf{w}) - \mathbf{x} - \nabla s, & q &= \nabla\cdot\mathbf{x}. \end{aligned}$$

Then p *is the solution of problem* (1.5), $A^{\frac{1}{2}}\mathbf{u}$ *is the flux,* $\mathbf{v} = \mathbf{0}$*, and* $q = 0$.

Proof. The proof follows immediately from Theorem 4.1. Let p satisfy (1.5) and let $\mathbf{u} = A^{\frac{1}{2}}\nabla p$, $\mathbf{v} = \mathbf{0}$, and $q = 0$. Then $\mathcal{L}_3(\mathbf{u}, p, \mathbf{v}, q)^t = (\mathbf{0}, f, \mathbf{0}, 0)^t$. Since both $\mathcal{L}_3$ and $\mathcal{L}_3^*$ are bijective, $(\mathbf{u}, p, \mathbf{v}, q)^t$ is unique and in the range of $\mathcal{L}_3^*$.

Coercivity of $\mathcal{L}_3^*$ implies that the solution, $(\mathbf{w}, r, \mathbf{x}, s)^t \in \mathcal{D}$, of (4.23) is the unique minimizer of $G_e(\mathbf{y}, \xi, \mathbf{z}, \eta) := \|\mathcal{L}_3^*(\mathbf{y}, \xi, \mathbf{z}, \eta)^t - (\mathbf{u}, p, \mathbf{v}, q)^t\|^2$ over $\mathcal{D}$, which implies that $\mathcal{L}_3^*(\mathbf{w}, r, \mathbf{x}, s)^t = (\mathbf{u}, p, \mathbf{v}, q)^t$. ∎

Remark 4.3. The solution, $(\mathbf{w}, r, \mathbf{x}, s)^t$, of (4.23) will have $s = 0$. Since the last two rows of $\mathcal{L}_3^*(\mathbf{w}, r, \mathbf{x}, s)^t = (\mathbf{u}, p, \mathbf{v}, q)^t$ are the same as the last two rows of (4.30), the same argument as in the proof above yields $s = 0$.

If, in addition, $\mathbf{b} = \mathbf{0}$, then $\mathbf{x} = \mathbf{0}$. To see this, apply $\nabla\times A^{-\frac{1}{2}}$ to the first row of $\mathcal{L}_3^*(\mathbf{w}, r, \mathbf{x}, 0)^t = (\mathbf{u}, p, \mathbf{v}, q)^t = (A^{\frac{1}{2}}\nabla p, p, \mathbf{0}, 0)^t$ and use the third row to get

$$\nabla\times A^{-\frac{1}{2}}\mathbf{w} + \nabla\times A^{-1}\nabla\times\mathbf{x} = \mathbf{x} + \nabla\times A^{-1}\nabla\times\mathbf{x} = \nabla\times\nabla p = \mathbf{0}.$$

Integration by parts yields

$$\langle\mathbf{x}, \mathbf{x}\rangle + \langle\nabla\times A^{-1}\nabla\times\mathbf{x}, \mathbf{x}\rangle = \langle\mathbf{x}, \mathbf{x}\rangle + \langle A^{-1}\nabla\times\mathbf{x}, \nabla\times\mathbf{x}\rangle = 0,$$

which implies $\mathbf{x} = \mathbf{0}$. A numerical scheme could make use of this information to simplify computations.

Again, consider the square system

$$\mathcal{L}_3\mathcal{L}_3^*(\mathbf{w}, r, \mathbf{x}, s)^t = (\mathbf{0}, -f, \mathbf{0}, 0)^t \quad \text{in } \Omega, \tag{4.32}$$

where

$$\mathcal{L}_3\mathcal{L}_3^* = ((\mathcal{L}_3\mathcal{L}_3^*)_1, (\mathcal{L}_3\mathcal{L}_3^*)_2, (\mathcal{L}_3\mathcal{L}_3^*)_3, (\mathcal{L}_3\mathcal{L}_3^*)_4)\,, \tag{4.33}$$

with

$$(\mathcal{L}_3\mathcal{L}_3^*)_1 = \begin{pmatrix} I - A^{\frac{1}{2}}\nabla\nabla\cdot A^{\frac{1}{2}} + A^{-\frac{1}{2}}\nabla\times\nabla\times A^{-\frac{1}{2}} \\ \nabla\cdot A^{\frac{1}{2}} - \mathbf{b}\cdot A^{-\frac{1}{2}} - c\nabla\cdot A^{\frac{1}{2}} \\ 0 \\ 0 \end{pmatrix},$$

$$(\mathcal{L}_3\mathcal{L}_3^*)_2 = \begin{pmatrix} -A^{\frac{1}{2}}\nabla - A^{-\frac{1}{2}}\mathbf{b} + A^{\frac{1}{2}}\nabla c \\ -\nabla\cdot A\nabla + \left(\mathbf{b}\cdot A^{-1}\mathbf{b} - (\nabla\cdot\mathbf{b}) + c^2\right) I \\ -\nabla\times A^{-1}\mathbf{b} \\ 0 \end{pmatrix},$$

$$(\mathcal{L}_3\mathcal{L}_3^*)_3 = \begin{pmatrix} 0 \\ -\mathbf{b}\cdot A^{-1}\nabla\times \\ I + \nabla\times A^{-1}\nabla\times - \nabla\nabla\cdot \\ -\nabla\cdot \end{pmatrix}, \quad (\mathcal{L}_3\mathcal{L}_3^*)_4 = \begin{pmatrix} 0 \\ 0 \\ \nabla \\ -\Delta \end{pmatrix}.$$

Notice that the diagonal of $\mathcal{L}_3\mathcal{L}_3^*$ consists of second-order operators, while the off-diagonal is first-order. Again, we say that $\mathcal{L}_3\mathcal{L}_3^*$ is *differentially diagonally dominant*. As in the two-dimensional case, this implies that convergence of the multilevel algorithm depends on $|\mathbf{b}|h$, where h is an appropriate finest-level mesh parameter, and the adverse effects of convection must diminish on finer meshes.

5. A FOSLL$^*_s$ two-stage approach. When there are no reaction terms in (1.5) (i.e., $c = 0$) and both Γ_D and Γ_N are connected, we can appeal to a FOSLL* *two-stage* scheme analogous to that for FOSLS [12, 8]. For focus here, we discuss the three-dimensional case, including comments on the two-dimensional case where appropriate.

The basic idea is that, in this case, slack variable $\mathbf{v}$ and the first and fourth equations in (4.16) (and the boundary conditions on p) are not needed to determine $\mathbf{u}$. (In two dimensions, the first equation in (4.1) would be omitted.) FOSLL* can thus be applied to the simpler first-stage system

$$\left\{ \begin{array}{rcll} -\nabla \cdot A^{\frac{1}{2}}\mathbf{u} + \mathbf{b} \cdot A^{-\frac{1}{2}}\mathbf{u} & = & f & \text{in} \quad \Omega, \\ \nabla \times A^{-\frac{1}{2}}\mathbf{u} - \nabla q & = & \mathbf{0} & \text{in} \quad \Omega \end{array} \right. \tag{5.1}$$

with boundary conditions

$$\left\{ \begin{array}{rcll} \mathbf{n} \cdot A^{\frac{1}{2}}\mathbf{u} & = & 0 & \text{on} \quad \Gamma_N, \\ \mathbf{n} \times A^{-\frac{1}{2}}\mathbf{u} & = & \mathbf{0} & \text{on} \quad \Gamma_D, \\ q & = & 0 & \text{on} \quad \Gamma_N. \end{array} \right. \tag{5.2}$$

(For the two-dimensional case, slack variable q in system (5.1), (5.2) is not needed.)

Changing the sign of the first equation in (5.1), the differential operator associated with this system is

$$\mathcal{L}_s = \begin{pmatrix} \nabla \cdot A^{\frac{1}{2}} - \mathbf{b} \cdot A^{-\frac{1}{2}} & 0 \\ \nabla \times A^{-\frac{1}{2}} & -\nabla \end{pmatrix} \tag{5.3}$$

with the formal adjoint

$$\mathcal{L}_s^* = \begin{pmatrix} -A^{\frac{1}{2}}\nabla - A^{-\frac{1}{2}}\mathbf{b} & A^{-\frac{1}{2}}\nabla \times \\ 0 & \nabla \cdot \end{pmatrix}. \tag{5.4}$$

Note that $\mathcal{L}_s$ may be obtained by eliminating rows 1 and 4 and columns 2 and 3 of $\mathcal{L}_3$. The differential operator for $\mathcal{L}_s^*$ can similarly be obtained from $\mathcal{L}_3^*$ by eliminating columns 1 and 4 and rows 2 and 3 of $\mathcal{L}_3^*$.

Here, $\mathcal{L}_s : \mathcal{D} \subseteq \mathcal{V} \to \mathcal{V}$, where $\mathcal{V} = L^2(\Omega)^{d+1}$ and $\mathcal{D} = \mathcal{W}_{ND}(A) \times H^1_N$. The domain of $\mathcal{L}_s^*$ involves the dual variables $(r, \mathbf{x})^t$ with $r \in H^1$ and $\mathbf{x} \in \mathcal{W}$. The boundary conditions associated with $\mathcal{L}_s^*$ can be determined by computing

$$\begin{aligned} \langle \mathcal{L}_s(\mathbf{u}, q)^t,\, (r, \mathbf{x})^t \rangle &= \langle (\mathbf{u}, q)^t,\, \mathcal{L}_s^*(r, \mathbf{x})^t \rangle \\ &\quad + \int_{\Gamma_D} (\mathbf{n} \cdot A^{\frac{1}{2}}\mathbf{u}) r - \int_{\Gamma_N} (\mathbf{n} \times A^{-\frac{1}{2}}\mathbf{u}) \cdot \mathbf{x} - \int_{\Gamma_D} q(\mathbf{n} \cdot \mathbf{x}), \end{aligned}$$

where (5.2) was used to reduce the boundary integral terms. The remaining boundary integrals vanish for every $(\mathbf{u}, q) \in \mathcal{D}$ if and only if $(r, \mathbf{x})$ satisfies

$$\left\{ \begin{array}{rcll} r & = & 0 & \text{on} \quad \Gamma_D, \\ \mathbf{n} \cdot \mathbf{x} & = & 0 & \text{on} \quad \Gamma_D, \\ \mathbf{n} \times \mathbf{x} & = & \mathbf{0} & \text{on} \quad \Gamma_N. \end{array} \right. \tag{5.5}$$

This yields $\mathcal{D}^* = H^1_N \times \mathcal{W}_{DN}$.

The associated bilinear forms are

$$
\begin{aligned}
a_s(\mathbf{u}, q;\, \mathbf{y}, \eta) &:= \langle \mathcal{L}_s(\mathbf{u}, q)^t,\, \mathcal{L}_s(\mathbf{y}, \eta)^t \rangle \\
&= \langle \nabla \cdot A^{\frac{1}{2}} \mathbf{u} - \mathbf{b} \cdot A^{-\frac{1}{2}} \mathbf{u},\, \nabla \cdot A^{\frac{1}{2}} \mathbf{y} - \mathbf{b} \cdot A^{-\frac{1}{2}} \mathbf{y} \rangle \\
&+ \langle \nabla \times A^{-\frac{1}{2}} \mathbf{u} - \nabla q,\, \nabla \times A^{-\frac{1}{2}} \mathbf{y} - \nabla \eta \rangle
\end{aligned} \tag{5.6}
$$

and

$$
\begin{aligned}
&a_s^*(r, \mathbf{x};\, \xi, \mathbf{z}) := \langle \mathcal{L}_3^*(r, \mathbf{x})^t,\, \mathcal{L}_s^*(\xi, \mathbf{z})^t \rangle \\
&\quad = \langle -A^{\frac{1}{2}} \nabla r - A^{-\frac{1}{2}} \mathbf{b} r + A^{-\frac{1}{2}} \nabla \times \mathbf{x},\, -A^{\frac{1}{2}} \nabla \xi - A^{-\frac{1}{2}} \mathbf{b} \xi + A^{-\frac{1}{2}} \nabla \times \mathbf{z} \rangle \\
&\quad + \langle \nabla \cdot \mathbf{x},\, \nabla \cdot \mathbf{z} \rangle,
\end{aligned} \tag{5.7}
$$

and the associated linear functional is

$$
f_s(\xi, \mathbf{z}) := \langle (-f, \mathbf{0})^t,\, (\xi, \mathbf{z})^t \rangle = -\langle f, \xi \rangle. \tag{5.8}
$$

The weak problem associated with the dual problem $\mathcal{L}_s^*(r, \mathbf{x})^t = (\mathbf{u}, q)^t$ is that of finding $(r, \mathbf{x}) \in \mathcal{D}^*$ such that

$$
a_s^*(r, \mathbf{x};\, \xi, \mathbf{z}) = f_s(\xi, \mathbf{z}) \tag{5.9}
$$

for every $(\xi, \mathbf{z}) \in \mathcal{D}^*$.

Note that

$$
a_s(\mathbf{u}, q;\, \mathbf{u}, q) + \langle \mathbf{u},\, \mathbf{u} \rangle = a_3(\mathbf{u}, 0, \mathbf{0}, q;\, \mathbf{u}, 0, \mathbf{0}, q), \tag{5.10}
$$

$$
a_s^*(r, \mathbf{x};\, r, \mathbf{x}) + \langle \mathbf{x},\, \mathbf{x} \rangle = a_3^*(\mathbf{0}, r, \mathbf{x}, 0;\, \mathbf{0}, r, \mathbf{x}, 0), \tag{5.11}
$$

where $a_3(\,\cdot\,;\,\cdot\,)$ and $a_3^*(\,\cdot\,;\,\cdot\,)$ are defined in (4.20), (4.21). We make use of this relationship in the following theorem.

THEOREM 5.1. *Operators $\mathcal{L}_s$ and $\mathcal{L}_s^*$ are bijective, and bilinear forms $a_s(\,\cdot\,;\,\cdot\,)$ and $a_s^*(\,\cdot\,;\,\cdot\,)$ are coercive and continuous on $\mathcal{D}$ and $\mathcal{D}^*$, respectively: There exist positive constants α_0, α_1 and α_0^*, α_1^*, which depend only on A, $\mathbf{b}$, c, and Ω, such that*

$$
\alpha_0 \, \|(\mathbf{u}, p)\|_{\mathcal{D}}^2 \le a_s(\mathbf{u}, p;\, \mathbf{u}, p) \le \alpha_1 \, \|(\mathbf{u}, p)\|_{\mathcal{D}}^2 , \tag{5.12}
$$

$$
\alpha_0^* \, \|(r, \mathbf{x})\|_{\mathcal{D}^*}^2 \le a_s^*(r, \mathbf{x};\, r, \mathbf{x}) \le \alpha_1^* \, \|(r, \mathbf{x})\|_{\mathcal{D}^*}^2 \tag{5.13}
$$

for any $(\mathbf{u}, p) \in \mathcal{D}$ and $(r, \mathbf{x}) \in \mathcal{D}^*$. Furthermore, problem (5.9) *has a unique solution $(r, \mathbf{x}) \in \mathcal{D}^*$ and it satisfies the a priori estimate*

$$
\|(r, \mathbf{x})\|_{\mathcal{D}^*} \le C \|f\|_{-1,D}. \tag{5.14}
$$

Proof. Continuity of the bilinear forms follows from repeated use of the triangle inequality. To establish coercivity of $a_s(\,\cdot\,;\,\cdot\,)$ and $a_s^*(\,\cdot\,;\,\cdot\,)$, we make use of Poincaré-type inequalities (see the discussion in [10]): When both Γ_D and Γ_N are connected, there exists a constant C, depending only on Ω, such that

$$
\begin{aligned}
\|\mathbf{u}\|^2 &\le C \left(\|\nabla \cdot A^{\frac{1}{2}} \mathbf{u}\|^2 + \|\nabla \times A^{-\frac{1}{2}} \mathbf{u}\|^2 \right), \\
\|\mathbf{x}\|^2 &\le C \left(\|\nabla \cdot \mathbf{x}\|^2 + \|\nabla \times \mathbf{x}\|^2 \right)
\end{aligned}
$$

for every $\mathbf{u} \in \mathcal{W}_{ND}(A)$ and $\mathbf{x} \in \mathcal{W}_{DN}$. These inequalities, together with a proof similar to that in Theorem 4.3, can be used to show that both $\mathcal{L}_s$ and $\mathcal{L}_s^*$ are injective. (See also the remark after Theorem 4.3.) Coercivity now follows from Theorem 4.3, observation (5.11), and a standard compactness argument (cf. [15]). We use Lemma 2.2 to establish that $\mathcal{L}$ and $\mathcal{L}^*$ are bijective.

It is easy to see that linear form $f_s(\cdot)$ is continuous in $\mathcal{D}^*$. Hence, by the Lax–Milgram lemma [15], problem (5.9) has a unique solution $(r, \mathbf{x}) \in \mathcal{D}^*$. Derivation of the a priori estimate is straightforward. ∎

We now establish that the solution of weak problem (5.9) yields the solution of system (1.5).

THEOREM 5.2. *Let* $(r, \mathbf{x}) \in \mathcal{D}^*$ *be the solution of* (5.9) *and let*

$$\mathbf{u} = -A^{\frac{1}{2}}\nabla r - A^{-\frac{1}{2}}\mathbf{b}r + A^{-\frac{1}{2}}\nabla\times\mathbf{x} \quad \textit{and} \quad v = \nabla\cdot\mathbf{x}.$$

Then $v = 0$ *and* $A^{\frac{1}{2}}\mathbf{u}$ *is the flux.*

Proof. The proof is the same as that of Theorem 4.4. ∎

As before, we examine the square system

$$\mathcal{L}_s\mathcal{L}_s^* = \begin{pmatrix} -\nabla\cdot A\nabla + \big(\mathbf{b}\cdot A^{-1}\mathbf{b} - (\nabla\cdot\mathbf{b})\big) I & -\mathbf{b}\cdot A^{-1}\nabla\times \\ -\nabla\times A^{-1}\mathbf{b} & \nabla\times A^{-1}\nabla\times - \nabla\nabla\cdot \end{pmatrix}. \tag{5.15}$$

This operator differs from the central 2×2 submatrix of $\mathcal{L}_3\mathcal{L}_3^*$ with $c = 0$ in that I is missing from its second diagonal term. The resulting second diagonal term in $\mathcal{L}_s\mathcal{L}_s^*$ is nonsingular only under the assumption that both Γ_D and Γ_N are connected. Both diagonal terms of $\mathcal{L}_s\mathcal{L}_s^*$ are Laplacian-like in that their weak forms are locally H^1-equivalent.

The diagonal part of $\mathcal{L}_s\mathcal{L}_s^*$ is dominant, especially for fine-grid scales, because the off-diagonal terms are first-order. This approach, therefore, retains the advantages of FOSLL$_e^*$ in that multigrid convergence should not degrade very quickly as the size of $\mathbf{b}$ increases, as the numerical results in Table 3 of section 7 confirm.

Original variable p can be recovered from a *second-stage* FOSLS scheme by minimizing $\|\nabla p - A^{-\frac{1}{2}}\mathbf{u}\|^2$ for this fixed $\mathbf{u}$. We call this combined approach FOSLL$_s^*$ for later reference. Finally, note that this approximation to p is obtained in the H^1 sense, which might be more desirable than the L^2 approximations obtained by the other FOSLL* schemes.

Remark 5.1. It is important to note that, in the definition of $\mathcal{L}_s^*$ (see (5.4)), the coefficient matrices appear outside of the differential operators. This implies that the discrete problem can be solved using finite-element spaces that do not have to take into account discontinuities of A. In two dimensions, domain $\mathcal{D}^* = H_N^1 \times H_D^1$, which implies that, even in the presence of re-entrant corners, finite-element spaces consisting of piecewise polynomials are sufficient. Of course, re-entrant corners do cause the discretization error to converge to zero at a reduced rate, as the results of section 7 show.

Remark 5.2. FOSLL$_s^*$ applies directly only to the case in which $c = 0$ and both Γ_D and Γ_N are connected. However, a simple modification to FOSLL$_s^*$ can be made to extend it to the more general case. Suppose that either Γ_D or Γ_N has multiple components. In two dimensions, $\mathcal{L}_s^*$ is still injective, but $\mathcal{L}_s$ now has a finite-dimensional null space. (For a description of the null space, see [10].) The unique solution, $(r, s)^t$, of the two-dimensional equivalent of (5.9) yields $\hat{\mathbf{u}} = \mathcal{L}_s^*(r, s)^t$, which differs from the desired solution, $\mathbf{u}$, by an element of the null space of $\mathcal{L}_s$. However, solution of the second stage,

$$p = \arg\min_{\hat{p}\in H_D^1} \|A^{-\frac{1}{2}}\hat{\mathbf{u}} - \nabla\hat{p}\|, \tag{5.16}$$

yields the solution of (1.5) because the null space of $\mathcal{L}_s$ is orthogonal to $\nabla\hat{p}$ for $\hat{p} \in H_D^1$. After p is computed, the flux can be computed as $A\nabla p$.

In three dimensions, both $\mathcal{L}_s$ and $\mathcal{L}_s^*$ have finite-dimensional null spaces. (For a description of the null spaces, see [10].) The solution of (5.9) is not unique, but any solution, $(r, \mathbf{x})^t$, yields $\hat{\mathbf{u}} = \mathcal{L}_s^*(r, \mathbf{x})^t$, which differs from the desired solution, $\mathbf{u}$, by an element of the null space of $\mathcal{L}_s$. Again, the solution of second-stage (5.16) yields the solution of (1.5), and the flux can be computed as $A\nabla p$.

Remark 5.3. Consider the case $c \neq 0$. A nonzero reaction term would be a subprincipal part of the operator, so one should be able to handle general case (1.5) in an iterative two-stage approach. This can be done in a unified way by defining (for the three-dimensional case) the abstract functional

$$\begin{aligned} F(\mathbf{u}, p, q;\, f) = \|&\mathbf{u} - A^{\frac{1}{2}}\nabla p\|^2 \\ &+ \|\mathcal{L}_s(\mathbf{u}, q)^t - (cp - f, \mathbf{0})^t\|_B^2, \end{aligned}$$

where the B-norm is defined in (1.1). Making the substitution $(\mathbf{u}, q)^t = \mathcal{L}_s^*(r, \mathbf{x})^t$, this becomes

$$\begin{aligned} F(p, r, \mathbf{x};\, f) = &\left\| \left(-A^{\frac{1}{2}}\nabla r - A^{-\frac{1}{2}}\mathbf{b}r + A^{-\frac{1}{2}}\nabla\times\mathbf{x} \right) - A^{\frac{1}{2}}\nabla p \right\|^2 \\ &+ \|\mathcal{L}_s\mathcal{L}_s^*(r, \mathbf{x})^t - (cp - f, \mathbf{0})^t\|_B^2. \end{aligned}$$

Note that the first term in F is just the usual FOSLS term $\|\mathbf{u} - A^{\frac{1}{2}}\nabla p\|^2$. The second term in F can be viewed as a first-stage FOSLS term for solving (5.1), where $-f$ on the right-hand side of the first equation is modified by adding cp.

We refer to F as an abstract functional because we do not expect to evaluate it in practice. To do so would require inversion of $\mathcal{L}_s\mathcal{L}_s^*$, that is, computation of $(\mathcal{L}_s\mathcal{L}_s^*)^{-1}(f - cp, \mathbf{0})^t$. The approach we suggest instead involves a block iteration process that alternates between $(r, \mathbf{x})^t$ and p updates. The iteration on $(r, \mathbf{x})^t$ would involve minimizing the second term in F with p fixed, which is equivalent to minimizing the functional

$$\|\mathcal{L}_s^*(r, \mathbf{x})^t\|^2 - 2\,\langle f - cp, r\rangle$$

over $(r, \mathbf{x})^t \in \mathcal{D}$. This avoids the need to invert $\mathcal{L}_s\mathcal{L}_s^*$. The p update would be done by minimizing the first term in F,

$$\left\| \left(-A^{\frac{1}{2}}\nabla r - A^{-\frac{1}{2}}\mathbf{b}r + A^{-\frac{1}{2}}\nabla\times\mathbf{x} \right) - A^{\frac{1}{2}}\nabla p \right\|^2,$$

with $(r, \mathbf{x})^t$ fixed.

6. Discrete approximation.

We first return to the abstract setting of section 2, where we introduced a generic operator $\mathcal{L} : \mathcal{D} \subseteq \mathcal{V}_1 \to \mathcal{V}_2$ and generic problem $\mathcal{L}\mathbf{u} = \mathbf{f}$. The assumptions on $\mathcal{L}$ and $\mathcal{D}$ imply that the adjoint operator, $\mathcal{L}^* : \mathcal{D}^* \subseteq \mathcal{V}_2 \to \mathcal{V}_1$, is well defined. If $\mathbf{u}$ is in the range of $\mathcal{L}^*$, say $\mathbf{u} = \mathcal{L}^*\mathbf{x}$, then we recast the problem as

$$\mathbf{x} = \arg\min_{\mathbf{y}\in\mathcal{D}^*} \|\mathcal{L}^*\mathbf{y} - \mathbf{u}\|_{\mathcal{V}_1}^2, \tag{6.1}$$

whose solution satisfies the weak problem of finding $\mathbf{x} \in \mathcal{D}^*$ such that

$$a^*(\mathbf{x};\, \mathbf{w}) := \langle \mathcal{L}^*\mathbf{x},\, \mathcal{L}^*\mathbf{w}\rangle_{\mathcal{V}_1} = \langle \mathbf{f},\, \mathbf{w}\rangle_{\mathcal{V}_2} \quad \forall \mathbf{w} \in \mathcal{D}^*. \tag{6.2}$$

Recall that if $\mathcal{D}^*$ is a Hilbert space that is continuously embedded in $\mathcal{V}_2$, and $\mathcal{L}^*$ has a bounded inverse, which implies that $a^*(\,\cdot\,;\,\cdot\,)$ is coercive on $\mathcal{D}^*$, then weak problem (6.2) is well posed and has a unique solution.

Here, we approximately minimize in (6.1) by restriction to a finite-dimensional subspace, say $\mathcal{V}_2^h \subset \mathcal{D}^*$. This yields the finite-dimensional weak problem of finding $\mathbf{x}^h \in \mathcal{V}_2^h$ such that

$$a^*(\mathbf{x}^h;\, \mathbf{w}^h) = \langle \mathbf{f},\, \mathbf{w}^h\rangle_{\mathcal{V}_2} \quad \forall \mathbf{w}^h \in \mathcal{V}_2^h. \tag{6.3}$$

Coercivity of $a^*(\,\cdot\,;\,\cdot\,)$ implies that (6.3) is well posed and has a unique solution.

Letting $\mathbf{u}^h = \mathcal{L}^*\mathbf{x}^h$, we clearly have

$$\|\mathbf{u} - \mathbf{u}^h\|_{\mathcal{V}_1} \le \inf_{\mathbf{y}^h \in \mathcal{V}_2^h} \|\mathcal{L}^*(\mathbf{x} - \mathbf{y}^h)\|_{\mathcal{V}_1} \le C \inf_{\mathbf{y}^h \in \mathcal{V}_2^h} \|\mathbf{x} - \mathbf{y}^h\|_{\mathcal{D}^*}. \tag{6.4}$$

In this paper, $\mathcal{V}_1$ and $\mathcal{V}_2$ are product $L^2(\Omega)$ spaces and $\mathcal{D}^*$ is locally H^1. We summarize the discussion in the following lemma.

LEMMA 6.1. *Let* $\mathbf{x} \in \mathcal{D}^*$ *and* $\mathbf{x}^h \in V_2^h$ *be the solutions of* (6.2) *and* (6.3), *respectively. Let* $\mathbf{u} = \mathcal{L}^*\mathbf{x}$ *and* $\mathbf{u}^h = \mathcal{L}^*\mathbf{x}^h$. *Then*

$$\|\mathbf{u} - \mathbf{u}^h\|_{\mathcal{V}_1} \le \inf_{\mathbf{y}^h \in \mathcal{V}_2^h} \|\mathcal{L}^*(\mathbf{x} - \mathbf{y}^h)\|_{\mathcal{V}_1} \le C \inf_{\mathbf{y}^h \in \mathcal{V}_2^h} \|\mathbf{x} - \mathbf{y}^h\|_{\mathcal{D}^*}. \tag{6.5}$$

Furthermore, if finite-dimensional subspace $\mathcal{V}_2^h$ *has the approximation property*

$$\inf_{\mathbf{y}^h \in \mathcal{V}_2^h} \|\mathbf{x} - \mathbf{y}^h\|_{\mathcal{D}^*} \le Ch^k, \tag{6.6}$$

then

$$\|\mathbf{u} - \mathbf{u}^h\|_{\mathcal{V}_1} \le Ch^k. \tag{6.7}$$

Proof. The proof is clear from the discussion above. □

The four adjoint operators, $\mathcal{L}_b^*, \mathcal{L}_2^*, \mathcal{L}_3^*$, and $\mathcal{L}_s^*$, defined in sections 3–5 are first-order, so that locally the desired approximation property (6.6) on finite-dimensional subspace $\mathcal{V}_2^h$ involves H^1-like norms. However, the $\mathcal{D}^*$-norm involves coefficient matrix A, which may be discontinuous. Assume that domain Ω is divided into a finite number of disjoint regions such that $\overline{\Omega} = \cup_{i=1}^p \overline{\Omega}_i$ and that coefficient matrix A is smooth in each region. Let $\mathcal{T}_h$ be a partition of Ω into finite elements: $\Omega = \cup_{K \in \mathcal{T}_h} K$ with $h = \max\{\mathrm{diam}(K) : K \in \mathcal{T}_h\}$. Assume that triangulation $\mathcal{T}_h$ is regular (see [15]) and that each K is completely contained in Ω_i for some i. Then it is possible to choose a finite-element basis that is conforming and satisfies the approximation property

$$\inf_{\mathbf{y}^h \in \mathcal{V}_2^h} \|\mathbf{x} - \mathbf{y}^h\|_{1,\,\Omega_i} \le Ch^k \|\mathbf{x}\|_{k+1,\,\Omega_i}, \tag{6.8}$$

where $k > 0$ is an integer and $\mathbf{x} \in H^{k+1}(\Omega_i)$ (cf. [18, 6]). A standard choice for $\mathcal{V}_2^h$ is the subspace of piecewise polynomials of degree k that are conforming and continuous on each Ω_i.

Define the *split norm*

$$\|\mathbf{w}\|_{k,S} = \left(\sum_{i=1}^{p} \|\mathbf{w}\|_{k,\Omega_i}^2 \right)^{\frac{1}{2}} \tag{6.9}$$

and associated space $H^k_S(\Omega)$. The following theorem provides a priori error estimates for each operator described above.

THEOREM 6.1. *We have the following estimates* (C *is a constant that depends only on* A, $\mathbf{b}$, c, *and* Ω)*:*

(1) *Assume that the solution,* $(\mathbf{w}, r, s)$, *of* (4.10) *is in* $H^{k+1}_S(\Omega)^4$, $(\mathbf{u}, p, q)$ *is defined according to Theorem* 4.2, *and* $(\mathbf{u}^h, p^h, q^h)$ *is the approximation defined in Lemma* 6.1. *Then*

$$\|A\nabla p - A^{\frac{1}{2}}\mathbf{u}^h\| + \|p - p^h\| + \|q^h\| \le Ch^k\|(\mathbf{w}, r, s)\|_{k+1,S}. \tag{6.10}$$

(2) *Assume that the solution,* $(\mathbf{w}, r, \mathbf{x}, s)$, *of* (4.23) *is in* $H^{k+1}_S(\Omega)^8$, $(\mathbf{u}, p, \mathbf{v}, q)$ *is defined according to Theorem* 4.4, *and* $(\mathbf{u}^h, p^h, \mathbf{v}^h, q^h)$ *is the approximation defined in Lemma* 6.1. *Then*

$$\|A\nabla p - A^{\frac{1}{2}}\mathbf{u}^h\| + \|p - p^h\| + \|\mathbf{v} - \mathbf{v}^h\| + \|q^h\| \le Ch^k\|(\mathbf{w}, r, \mathbf{x}, s)\|_{k+1,S}. \tag{6.11}$$

(3) *Assume that the solution,* $(r, \mathbf{x})$, *of* (5.9) *is in* $H^{k+1}_S(\Omega)^{d+1}$, $(\mathbf{u}, q)$ *is defined according to Theorem* 5.2, *and* $(\mathbf{u}^h, q^h)$ *is the approximation defined in Lemma* 6.1. *Then*

$$\|A^{\frac{1}{2}}\mathbf{u} - A^{\frac{1}{2}}\mathbf{u}^h\| + \|q^h\| \le Ch^k\|(r, \mathbf{x})\|_{k+1,S}. \tag{6.12}$$

Proof. The proof is a simple consequence of Theorems 4.2, 4.4, and 5.2 and Lemma 6.1. ∎

Remark 6.1. Under the same stronger smoothness assumptions on coefficient matrix A and domain Ω as those in [10], one can show that

$$\frac{1}{C}\|\mathbf{v}\|_1^2 \le \|\mathbf{v}\|^2 + \|\nabla \cdot A^{\frac{1}{2}}\mathbf{v}\|^2 + \|\nabla \times A^{-\frac{1}{2}}\mathbf{v}\|^2 \le C\,\|\mathbf{v}\|_1^2$$

for $\mathbf{v}$ in the spaces we consider in sections 3–5. This equivalence estimate and Theorems 4.1, 4.3, and 5.1 confirm that multiplicative and additive multigrid algorithms applied to the resulting discrete problems for the dual variable are optimally convergent (see [10]).

7. Numerical results. We tested FOSLL$^*_e$ and FOSLL$^*_s$ on problem (1.5) in two dimensions. In all but the last two sets of tests, we chose A to be the identity matrix, the domain to be the unit square, and the exact solution to be $p = x(x-1)\sin(\pi y)$. Several choices for $\mathbf{b}$ and c were used, and the exact values of the other unknowns and the right-hand sides were defined consistently. The minimization problems were discretized using bilinear finite elements on a uniform square mesh, with h ranging from $\frac{1}{4}$ to $\frac{1}{128}$.

For simplicity, we used algebraic multigrid (AMG) [5] to solve the discrete systems, with separate coarse grids and intergrid transfer operators determined for each unknown. (This is not a block relaxation scheme because, while coarsening is set up within each level, all variables are coupled in the coarse-grid correction process.) We started with (1,1) V-cycles based on the so-called C/F Gauss–Seidel relaxation scheme. When required, a normalization step was used after each fine-grid relaxation sweep to eliminate arbitrary constants in the approximations. The results presented include asymptotic convergence factors per cycle for AMG (ρ); the L^2 error for p, (ϵ^0_p); the H^1 error for p, (ϵ^1_p), where appropriate; and the L^2 error for $\mathbf{u}$, ($\epsilon_{\mathbf{u}}$).

Table 1
Results of V*-cycle tests for FOSLL$^*_e$.*

$\mathbf{b}^t$	c	h	ρ	ϵ_p^0	$\epsilon_{\mathbf{u}}$
$(0,0)$	0	$\frac{1}{4}$	0.06	$2.25e-2$	$1.43e-1$
		$\frac{1}{8}$	0.09	$1.07e-2$	$7.18e-2$
		$\frac{1}{16}$	0.10	$5.26e-3$	$3.59e-2$
		$\frac{1}{32}$	0.10	$2.62e-3$	$1.80e-2$
		$\frac{1}{64}$	0.10	$1.31e-3$	$8.98e-3$
		$\frac{1}{128}$	0.10	$6.54e-4$	$4.49e-3$
$(2,3)$	0	$\frac{1}{4}$	0.18	$2.60e-2$	$1.89e-1$
		$\frac{1}{8}$	0.24	$5.33e-3$	$9.64e-2$
		$\frac{1}{16}$	0.25	$1.12e-2$	$4.84e-2$
		$\frac{1}{32}$	0.26	$2.63e-3$	$2.43e-2$
		$\frac{1}{64}$	0.26	$1.31e-3$	$1.21e-2$
		$\frac{1}{128}$	0.26	$6.54e-4$	$6.07e-3$
$(4,6)$	0	$\frac{1}{4}$	0.39	$3.47e-2$	$2.58e-1$
		$\frac{1}{8}$	0.54	$1.31e-2$	$1.37e-1$
		$\frac{1}{16}$	0.56	$5.62e-3$	$6.95e-2$
		$\frac{1}{32}$	0.56	$2.67e-3$	$3.49e-2$
		$\frac{1}{64}$	0.56	$1.32e-3$	$1.75e-2$
		$\frac{1}{128}$	0.56	$6.55e-4$	$8.73e-3$
$(6,9)$	0	$\frac{1}{4}$	0.33	$4.38e-2$	$3.08e-1$
		$\frac{1}{8}$	0.71	$1.59e-2$	$1.71e-1$
		$\frac{1}{16}$	0.71	$6.15e-3$	$8.84e-2$
		$\frac{1}{32}$	0.73	$2.74e-3$	$4.46e-2$
		$\frac{1}{64}$	0.73	$1.32e-3$	$2.24e-2$
		$\frac{1}{128}$	0.73	$6.56e-4$	$1.12e-2$
$(0,0)$	-1	$\frac{1}{4}$	0.06	$7.10e-3$	$1.37e-1$
		$\frac{1}{8}$	0.09	$1.76e-3$	$6.87e-2$
		$\frac{1}{16}$	0.09	$4.38e-4$	$3.43e-2$
		$\frac{1}{32}$	0.10	$1.09e-4$	$1.72e-2$
		$\frac{1}{64}$	0.10	$2.73e-5$	$8.59e-3$
		$\frac{1}{128}$	0.10	$6.84e-6$	$4.29e-3$
$(0,0)$	-10	$\frac{1}{4}$	0.21	$1.15e-1$	$1.16e-1$
		$\frac{1}{8}$	0.50	$6.17e-2$	$5.23e-2$
		$\frac{1}{16}$	0.51	$3.14e-2$	$2.53e-2$
		$\frac{1}{32}$	0.52	$1.58e-2$	$1.26e-2$
		$\frac{1}{64}$	0.52	$7.90e-3$	$6.26e-3$
		$\frac{1}{128}$	0.52	$3.95e-3$	$3.13e-3$

Table 1 contains results obtained for FOSLL$^*_e$. Errors in p and $\mathbf{u}$ were measured as follows. Let p and $\mathbf{u}$ denote the exact solutions to the continuous first-order problem. Define

$$\epsilon_p^0 = \|p - p^h\|\,, \quad \text{where } p^h = \nabla\cdot A^{\frac{1}{2}}\mathbf{w}^h - cr^h\,,$$

and

$$\epsilon_{\mathbf{u}} = \|\mathbf{u} - \mathbf{u}^h\|\,, \quad \text{where } \mathbf{u}^h = \mathbf{w}^h - A^{\frac{1}{2}}\nabla r^h - A^{-\frac{1}{2}}\mathbf{b}r^h + A^{-\frac{1}{2}}\nabla^\perp s^h\,.$$

Generally, the method behaved as expected, with apparent $O(h)$ approximation errors. An exception is the case $\mathbf{b} = (0,0)^t, c = -1$, where ϵ_p^0 appears to be $O(h^2)$.

Table 2
Results of V*-cycle tests for FOSLL$^*_s$.*

$\mathbf{b}^t$	c	h	ρ	ϵ_p^0	ϵ_p^1	$\epsilon_{\mathbf{u}}$
$(0,0)$	0	$\frac{1}{4}$	0.05	$7.35e-3$	$1.37e-1$	$1.37e-1$
		$\frac{1}{8}$	0.06	$1.83e-3$	$6.87e-2$	$6.87e-2$
		$\frac{1}{16}$	0.09	$4.56e-4$	$3.43e-2$	$3.43e-2$
		$\frac{1}{32}$	0.12	$1.14e-4$	$1.72e-2$	$1.72e-2$
		$\frac{1}{64}$	0.15	$2.85e-5$	$8.59e-3$	$8.59e-3$
		$\frac{1}{128}$	0.16	$7.13e-6$	$4.29e-3$	$4.29e-3$
$(2,3)$	0	$\frac{1}{4}$	0.17	$1.29e-2$	$1.40e-1$	$1.82e-1$
		$\frac{1}{8}$	0.30	$3.34e-3$	$6.91e-2$	$9.22e-2$
		$\frac{1}{16}$	0.30	$8.42e-4$	$3.44e-2$	$4.63e-2$
		$\frac{1}{32}$	0.29	$2.11e-4$	$1.72e-2$	$2.32e-2$
		$\frac{1}{64}$	0.30	$5.28e-5$	$8.59e-3$	$1.16e-2$
		$\frac{1}{128}$	0.30	$1.32e-5$	$4.29e-3$	$5.79e-3$
$(4,6)$	0	$\frac{1}{4}$	0.39	$2.48e-2$	$1.61e-1$	$2.50e-1$
		$\frac{1}{8}$	0.55	$6.86e-3$	$7.28e-2$	$1.31e-1$
		$\frac{1}{16}$	0.56	$1.77e-3$	$3.49e-2$	$6.64e-2$
		$\frac{1}{32}$	0.55	$4.45e-4$	$1.73e-2$	$3.33e-2$
		$\frac{1}{64}$	0.56	$1.11e-4$	$8.60e-3$	$1.67e-2$
		$\frac{1}{128}$	0.57	$2.79e-5$	$4.29e-3$	$8.33e-3$
$(6,9)$	0	$\frac{1}{4}$	0.34	$3.59e-2$	$1.95e-1$	$2.99e-1$
		$\frac{1}{8}$	0.70	$1.09e-2$	$8.19e-2$	$1.64e-1$
		$\frac{1}{16}$	0.71	$2.89e-3$	$3.64e-2$	$8.45e-2$
		$\frac{1}{32}$	0.71	$7.34e-4$	$1.75e-2$	$4.26e-2$
		$\frac{1}{64}$	0.71	$1.84e-4$	$8.62e-3$	$2.13e-2$
		$\frac{1}{128}$	0.72	$4.61e-5$	$4.30e-3$	$1.07e-2$

Although this deserves a closer look, it is probably due to a lucky choice of the combination of exact solution and coefficients. When convection is present, the errors in p are apparently $O(h)$. AMG convergence is affected by increasingly larger convection. There appears to be sensitivity to the presence of reaction terms ($c \neq 0$), although this sensitivity is not as dramatic as with convection.

Table 2 contains results for FOSLL$^*_s$ for the test problems without reaction ($c = 0$). Since p^h is obtained in the second stage of the solution method, it is not necessary here to reconstruct it as we did before. Also, we are now able to control H^1 accuracy in p, so we define the following error measures:

$$\epsilon_p^0 = \|p - p^h\| \quad (L^2\,\text{error})\,,$$

$$\epsilon_p^1 = \|\mathbf{u} - \nabla p^h\| \quad (H^1\,\text{error})\,,$$

and

$$\epsilon_{\mathbf{u}} = \|\mathbf{u} - \mathbf{u}^h\|\,, \quad \text{where } \mathbf{u}^h = \mathbf{w}^h - A^{\frac{1}{2}}\nabla r^h - A^{-\frac{1}{2}}\mathbf{b}r^h + A^{-\frac{1}{2}}\nabla^\perp s^h\,.$$

AMG convergence is very close to that obtained for FOSLL$^*_e$. Discretization errors ϵ_p^1 and $\epsilon_{\mathbf{u}}$ both behave like $O(h)$, while the L^2 error ϵ_p^0 behaves like $O(h^2)$.

Our V-cycle results are sensitive to the size of convection. For both FOSLL$^*_e$ and FOSLL$^*_s$, the off-diagonal coupling in the formal normal is first-order, which

TABLE 3
Results of W*-cycle tests for both methods.*

$\mathbf{b}^t$	c	h	FOSLL$^*_e$	FOSLL$^*_s$
$(6,9)$	0	$\frac{1}{4}$	0.33	0.34
		$\frac{1}{8}$	0.64	0.64
		$\frac{1}{16}$	0.51	0.51
		$\frac{1}{32}$	0.30	0.31
		$\frac{1}{64}$	0.11	0.15
		$\frac{1}{128}$	0.04	0.08

means that the degradation of convergence tends to happen only at relatively coarse scales. This reasoning suggests that W-cycles, which are still of optimal cost, would be more effective. This intuition is confirmed by the (1,1) W-cycle results for the case $\mathbf{b}^t = (6,9), c = 0$ shown in Table 3. Note the improvement for both FOSLL$^*_e$ and FOSLL$^*_s$ as h decreases. This indicates that multigrid performance for these methods is dictated primarily by $|\mathbf{b}|h$, where h is the fine-grid mesh size. Note also that these methods are not limited by a sufficiently fine coarsest grid, as are many standard multigrid methods for convection-diffusion problems.

Next, we present results for the W-cycle version of FOSLL$^*_s$ applied to a discontinuous coefficient problem. We again treat (1.5), but now with $A = aI$, where I is the 2×2 identity matrix and a is the function defined for a given constant σ on the unit square by

$$a(x,y) = \begin{cases} 1, & x \le \frac{1}{2}, \\ \sigma, & x > \frac{1}{2}. \end{cases} \tag{7.1}$$

We constructed exact solution p so that $\mathbf{u}$ is not a product H^1 function:

$$p(x,y) = \begin{cases} \big((2\sigma-4)*x^2 + (4-\sigma)*x\big)\sin(\pi y), & x \le \frac{1}{2}, \\ \big(-6*x^2 + 7*x - 1\big)\sin(\pi y), & x > \frac{1}{2}. \end{cases} \tag{7.2}$$

Table 4 shows results for the cases $\mathbf{b}^t = (0,0), c = 0$ and $\mathbf{b}^t = (6,9), c = 0$ with the three values $\sigma = 1, 10, 100$. As expected, the convergence factors depend on the size of σ, but the degradation is fairly graceful, an indication of the robustness of the method.

We end this paper with results on FOSLL$^*_s$ for a re-entrant corner problem with a solution that involves the singular function. We return to the case $A = I$, but now with the domain being the unit square where the lower right quadrant is removed, producing a re-entrant corner at $(0.5, 0.5)$. Letting (r,θ) be the polar coordinate system centered at $(0.5, 0.5)$, define $p = \delta(r) r^{2/3} \sin(2\pi\theta/3)$, where $\delta(r)$ is the C^2 *cut-off* function defined by the properties $\delta(r) = 1$ for $r < 1/8$, $\delta(r) = 0$ for $r > 3/8$, and $\delta(r)$ is a quintic polynomial in $[1/8, 3/8]$. This cut-off function is used to give homogeneous boundary conditions for p, which is unnecessary for our method but convenient for our code. We used W-cycles here for all tests. Table 5 shows results for the cases $\mathbf{b}^t = (0,0), c = 0$ and $\mathbf{b}^t = (6,9), c = 0$. AMG convergence is similar to that obtained in Table 4 for the corresponding problems without the re-entrant corner. With and without convection, error ϵ^0_p behaves like $O(h^{5/3})$, which is to be expected because $p \in H^{5/3}$. The H^1 errors in p and L^2 errors in $\mathbf{u}$ both behave a little worse than $O(h)$, but a little better than the expected $O(h^{2/3})$.

TABLE 4
Results of W*-cycle tests for* $FOSLL_s^*$ *with coefficient jump factors of* $\sigma = 1, 10, 100$.

σ	$\mathbf{b}^t$	c	h	ρ	ϵ_p^0	ϵ_p^1	$\epsilon_{\mathbf{u}}$
1	$(0,0)$	0	$\frac{1}{4}$	0.05	$3.23e-2$	$5.88e-1$	$5.88e-1$
			$\frac{1}{8}$	0.05	$8.00e-3$	$2.94e-1$	$2.94e-1$
			$\frac{1}{16}$	0.07	$2.00e-4$	$1.47e-1$	$1.47e-1$
			$\frac{1}{32}$	0.07	$5.00e-4$	$7.35e-2$	$7.35e-2$
			$\frac{1}{64}$	0.07	$1.25e-4$	$3.68e-2$	$3.68e-2$
			$\frac{1}{128}$	0.07	$3.12e-5$	$1.84e-2$	$1.84e-2$
10	$(0,0)$	0	$\frac{1}{4}$	0.07	$7.91e-2$	$1.30e+0$	$2.03e+0$
			$\frac{1}{8}$	0.12	$1.93e-2$	$6.41e-1$	$1.01e+0$
			$\frac{1}{16}$	0.13	$4.80e-3$	$3.20e-1$	$5.05e-1$
			$\frac{1}{32}$	0.12	$1.20e-3$	$1.60e-1$	$2.52e-1$
			$\frac{1}{64}$	0.12	$3.00e-4$	$7.98e-2$	$1.26e-1$
			$\frac{1}{128}$	0.12	$7.49e-5$	$3.99e-2$	$6.30e-2$
100	$(0,0)$	0	$\frac{1}{4}$	0.15	$9.18e-1$	$1.47e+1$	$1.56e+1$
			$\frac{1}{8}$	0.23	$2.25e-1$	$7.26e+0$	$7.71e+0$
			$\frac{1}{16}$	0.26	$5.59e-2$	$3.62e+0$	$3.85e+0$
			$\frac{1}{32}$	0.24	$1.40e-2$	$1.81e+0$	$1.92e+0$
			$\frac{1}{64}$	0.26	$3.49e-3$	$9.05e-1$	$9.61e-1$
			$\frac{1}{128}$	0.24	$8.72e-4$	$4.52e-1$	$4.80e-1$
1	$(6,9)$	0	$\frac{1}{4}$	0.34	$1.51e-1$	$8.29e-1$	$1.27e-1$
			$\frac{1}{8}$	0.64	$4.55e-2$	$3.49e-1$	$6.94e-1$
			$\frac{1}{16}$	0.51	$1.21e-2$	$1.56e-1$	$3.57e-1$
			$\frac{1}{32}$	0.31	$3.07e-3$	$7.47e-2$	$1.80e-1$
			$\frac{1}{64}$	0.15	$7.70e-4$	$3.69e-2$	$9.02e-1$
			$\frac{1}{128}$	0.08	$1.93e-4$	$1.84e-2$	$4.51e-2$
10	$(6,9)$	0	$\frac{1}{4}$	0.36	$1.45e-1$	$1.50e+0$	$2.87e+0$
			$\frac{1}{8}$	0.49	$4.40e-2$	$6.89e-1$	$1.51e+0$
			$\frac{1}{16}$	0.43	$1.17e-2$	$3.27e-1$	$7.69e-1$
			$\frac{1}{32}$	0.27	$2.96e-3$	$1.61e-1$	$3.87e-1$
			$\frac{1}{64}$	0.16	$7.43e-4$	$7.99e-2$	$1.94e-1$
			$\frac{1}{128}$	0.13	$1.86e-4$	$3.99e-2$	$9.68e-2$
100	$(6,9)$	0	$\frac{1}{4}$	0.33	$2.01e+0$	$1.72e+1$	$2.21e+1$
			$\frac{1}{8}$	0.48	$6.44e-1$	$7.92e+0$	$1.23e+1$
			$\frac{1}{16}$	0.44	$1.74e-1$	$3.73e+0$	$6.38e+0$
			$\frac{1}{32}$	0.29	$4.45e-2$	$1.82e+0$	$3.22e+0$
			$\frac{1}{64}$	0.26	$1.12e-2$	$9.07e-1$	$1.61e+0$
			$\frac{1}{128}$	0.25	$2.80e-3$	$4.53e-1$	$8.07e-1$

When $\mathbf{b}^t = (0,0)$, the curl equation in (5.1) is not really needed. Eliminating this equation amounts to dropping the second row of (5.3) and second column of (5.4), which is the same as setting $\mathbf{x} = \mathbf{z} = \mathbf{0}$ in (5.9). The resulting FOSLL* scheme actually reduces to the usual Galerkin method for solving (1.5) for this special case. The results obtained with this simpler scheme are similar to what is shown in Table 5 for p. The benefits of our FOSLL* techniques are thus evident in their ability to handle general diffusion-convection-reaction problems in a fully variational framework with Poisson-like efficiency and accuracy.

Table 5
Results of W*-cycle tests for FOSLL$^*_s$ with a re-entrant corner.*

$\mathbf{b}^t$	c	h	ρ	ϵ_p^0	ϵ_p^1	$\epsilon_{\mathbf{u}}$
$(0,0)$	0	$\frac{1}{4}$	0.05	$4.22e-2$	$6.33e-1$	$6.68e-1$
		$\frac{1}{8}$	0.05	$1.22e-2$	$4.12e-1$	$4.18e-1$
		$\frac{1}{16}$	0.05	$2.94e-3$	$1.86e-1$	$1.87e-1$
		$\frac{1}{32}$	0.05	$8.14e-4$	$9.63e-2$	$9.70e-2$
		$\frac{1}{64}$	0.05	$2.33e-4$	$5.01e-2$	$5.06e-2$
		$\frac{1}{128}$	0.05	$7.12e-5$	$2.64e-2$	$2.68e-2$
$(6,9)$	0	$\frac{1}{4}$	0.32	$4.72e-2$	$6.56e-1$	$6.02e-1$
		$\frac{1}{8}$	0.46	$1.57e-2$	$4.22e-1$	$4.14e-1$
		$\frac{1}{16}$	0.38	$4.05e-3$	$1.87e-1$	$1.90e-1$
		$\frac{1}{32}$	0.16	$1.10e-3$	$9.65e-2$	$9.72e-2$
		$\frac{1}{64}$	0.05	$3.06e-4$	$5.01e-2$	$4.95e-2$
		$\frac{1}{128}$	0.05	$8.92e-5$	$2.64e-2$	$2.53e-2$

REFERENCES

[1] P. B. Bochev, *Analysis of least-squares finite element methods for the Navier–Stokes equations*, SIAM J. Numer. Anal., 34 (1997), pp. 1817–1844.

[2] P. Bochev, Z. Cai, T. A. Manteuffel, and S. F. McCormick, *Analysis of velocity-flux first-order system least-squares principles for the Navier–Stokes equations: Part* I, SIAM J. Numer. Anal., 35 (1998), pp. 990–1009.

[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] J. H. Bramble, R. D. Lazarov, and J. E. Pasciak, *A least-squares approach based on a discrete minus one inner product for first order system*, Math. Comp., 66 (1997), pp. 935–955.

[5] A. Brandt, S. McCormick, and J. Ruge, *Algebraic Multigrid (AMG) for Automatic Multigrid Solution with Application to Geodetic Computations*, Tech. report, Inst. Comp. Studies, Colorado State University, Boulder, CO, 1982.

[6] M. Berndt, *Adaptive Refinement and the Treatment of Discontinuous Coefficients for Multilevel First-Order System Least Squares (FOSLS)*, Ph.D. thesis, University of Colorado at Boulder, Boulder, CO, 1999.

[7] 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.

[8] Z. Cai, C.-O. Lee, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for linear elasticity equations: Further results*, SIAM J. Sci. Comput., 21 (2000), pp. 1728–1739.

[9] Z. Cai, T. Manteuffel, and S. McCormick, *First-order system least squares for velocity-vorticity-pressure form of the Stokes equations, with application to linear elasticity*, Electron. Trans. Numer. Anal., 3 (1995), pp. 150–159.

[10] 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.

[11] Z. Cai, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for the Stokes equations, with application to linear elasticity*, SIAM J. Numer. Anal., 34 (1997), pp. 1727–1741.

[12] Z. Cai, T. A. Manteuffel, S. F. McCormick, and S. V. Parter, *First-order system least squares (FOSLS) for planar linear elasticity: Pure traction problem*, SIAM J. Numer. Anal., 35 (1998), pp. 320–335.

[13] C. L. Chang, *Finite element approximation for grad-div type systems in the plane*, SIAM J. Numer. Anal., 29 (1992), pp. 452–461.

[14] T. F. Chen and G. J. Fix, *Least-squares finite element simulation of transonic flows*, Appl. Numer. Math., 2 (1986), pp. 399–408.

[15] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, New York, 1978.

[16] 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., 19 (1998), pp. 1958–1979.

[17] V. Girault and P. A. Raviart, *Finite Element Methods for Navier–Stokes Equations: Theory and Algorithms*, Springer-Verlag, New York, 1986.

[18] T. A. Manteuffel, S. F. McCormick, and G. Starke, *First-order system least squares for second-order problems with discontinuous coefficients*, in Proceedings of the Seventh Copper Mountain Conference on Multigrid Methods, NASA Publications, Hampton, VA, 1995, pp. 535–550.

[19] A. I. Pehlivanov, G. F. Carey, and R. D. Lazarov, *Least-squares mixed finite elements for second-order elliptic problems*, SIAM J. Numer. Anal., 31 (1994), pp. 1368–1377.

[20] K. Tanabe, *Projection methods for solving singular systems of linear equations and its applications*, Numer. Math., 17 (1971), pp. 203–214.