SIAM J. SCI. COMPUT.
Vol. 26, No. 1, pp. 31–54

LEAST-SQUARES FINITE ELEMENT METHODS AND ALGEBRAIC MULTIGRID SOLVERS FOR LINEAR HYPERBOLIC PDEs*

H. DE STERCK†, THOMAS A. MANTEUFFEL†, STEPHEN F. MCCORMICK†, AND LUKE OLSON‡

Abstract. Least-squares finite element methods (LSFEMs) for scalar linear partial differential equations (PDEs) of hyperbolic type are studied. The space of admissible boundary data is identified precisely, and a trace theorem and a Poincaré inequality are formulated. The PDE is restated as the minimization of a least-squares functional, and the well-posedness of the associated weak formulation is proved. Finite element convergence is proved for conforming and nonconforming (discontinuous) LSFEMs that are similar to previously proposed methods but for which no rigorous convergence proofs have been given in the literature. Convergence properties and solution quality for discontinuous solutions are investigated in detail for finite elements of increasing polynomial degree on triangular and quadrilateral meshes and for the general case that the discontinuity is not aligned with the computational mesh. Our numerical studies found that higher-order elements yield slightly better convergence properties when measured in terms of the number of degrees of freedom. Standard algebraic multigrid methods that are known to be optimal for large classes of elliptic PDEs are applied without modifications to the linear systems that result from the hyperbolic LSFEM formulations. They are found to yield complexity that grows only slowly relative to the size of the linear systems.

Key words. least-squares variational formulation, finite element discretization, hyperbolic problems, algebraic multigrid

AMS subject classifications. 65N15, 65N30, 65N55

DOI. 10.1137/S106482750240858X

1. Introduction.

We consider scalar linear partial differential equations (PDEs) of hyperbolic type that are of the form

$$\mathbf{b} \cdot \nabla p = f \quad \text{in } \Omega, \tag{1.1}$$

$$p = g \quad \text{on } \Gamma_I, \tag{1.2}$$

with $\mathbf{b}(\mathbf{x})$ a flow field on $\Omega \subset \mathbb{R}^d$, and

$$\Gamma_I := \{\mathbf{x} \in \partial\Omega \colon \mathbf{n}(\mathbf{x}) \cdot \mathbf{b}(\mathbf{x}) < 0\}, \tag{1.3}$$

the inflow part of the boundary of domain Ω. Here, $\mathbf{n}(\mathbf{x})$ is the outward unit normal of $\partial\Omega$.

Equations of this type, often called transport equations or linear advection equations, arise in many applications in science and engineering, e.g., in fluid dynamics [22] and in neutron transport [23]. For decades there has been a drive to find increasingly accurate and efficient numerical solution methods for equations of the form (1.1)–(1.2). Not only do these equations have wide applications by themselves, but they also form a prototype equation for more general equations of hyperbolic type,

*Received by the editors June 1, 2002; accepted for publication (in revised form) October 23, 2003; published electronically August 19, 2004. This work was sponsored by the Department of Energy under grants DE-FC02-01ER25479 and DE-FG02-03ER25574, Lawrence Livermore National Laboratory under contract B533502, Sandia National Laboratory under contract 15268, and the National Science Foundation under VIGRE grant DMS-9810751.

http://www.siam.org/journals/sisc/26-1/40858.html

†Department of Applied Mathematics, Campus Box 526, University of Colorado at Boulder, Boulder, CO 80302 (desterck@colorado.edu, tmanteuf@colorado.edu, stevem@colorado.edu).

‡Division of Applied Mathematics, Brown University, 182 George Street, Box F, Providence, RI 02912 (lolson@dam.brown.edu).

e.g., systems of nonlinear conservation laws [22] or transport equations in phase space [23]. Successful numerical methods for (1.1)–(1.2) can often be used as building blocks for the numerical solution of more complicated hyperbolic PDEs [22].

Linear hyperbolic PDEs allow for discontinuous solutions when the boundary data is discontinuous. It is difficult to develop numerical methods that offer both high-order accurate results in regions of smooth solution and sharp discontinuity resolution, while avoiding spurious oscillations at discontinuities [3]. For wide classes of elliptic PDEs, optimal multilevel iterative solution algorithms have been developed for the discrete linear algebraic systems that require only $O(n)$ operations, where n is the number of unknowns (see, e.g., [29, 10] and references therein). For hyperbolic and mixed elliptic-hyperbolic PDEs, attempts at finding such optimal iterative solvers have been scarcely successful, even though some promising results have been reported [30].

The general philosophy behind the approach pursued in this paper is to combine adaptive least-squares (LS) finite element discretizations on space-time domains with global implicit solves using optimal iterative methods, in particular algebraic multigrid (AMG). Our goal is to explore whether such an approach can be competitive with present-day state-of-the-art techniques, e.g., approaches that rely on explicit time-marching using discontinuous Galerkin (DG) schemes. Clearly, there are important difficulties that have to be overcome. Optimal $O(n)$ solvers are still an active research topic for general hyperbolic and mixed elliptic-hyperbolic PDEs. A strong motivation for our choice of LS discretizations is that optimal solvers are more easily designed for the symmetric positive-definite (SPD) matrices that result from LS discretizations. We intend to remedy the extra smearing at discontinuities that is introduced by LS methods, as compared with other approaches, by adaptive refinement based on the natural, sharp error estimator provided by the LS functional; see Remark 3.7. The research question we seek to answer is whether the resulting adaptive least-squares finite element methods (LSFEMs), combined with optimal solvers, can be competitive with other approaches. The scope of the present paper encompasses the theoretical aspects of the LSFEM for continuous and for discontinuous elements, as well as a numerical study of AMG performance. Work on the combination of these techniques with adaptive refinement is the subject of a forthcoming paper. Application of these methods to linear hyperbolic systems and general systems of nonlinear conservation laws is also work in progress.

FEMs for (1.1)–(1.2) have been considered before, for example in Galerkin, streamline-upwind Petrov–Galerkin (SUPG) and residual distribution frameworks [20, 15, 1]. LS terms have been added to Galerkin methods for stabilization (see e.g. [17, 2]), and the SUPG method can be written as a linear combination of a Galerkin method and a LS term [15]. A comparison of Galerkin, SUPG, and LSFEM for convection problems can be found in [5]. In the present paper, we investigate pure LS formulations for (1.1)–(1.2). While LSFEMs have been investigated extensively for equations of elliptic type [11, 12, 19, 7], their use for hyperbolic PDEs has been initiated only recently [13, 6, 17]. LSFEMs are inherently attractive variational formulations for which well-posedness of the resulting discrete problems can be proved rigorously. LS finite element formulations lead to SPD linear systems. Another advantage of the LSFEM approach is that higher-order accurate methods can easily be constructed which are linear (for linear or linearized PDEs). As shown in this paper, these linear higher-order discretizations do not exhibit excessive spurious oscillations at discontinuities. This is in contrast to most other methods, e.g., DG methods, where nonlinear limiter functions have to be employed in order to assure

monotonicity [14]. Linear discretizations are better suited for iterative solution in global implicit solves. In the case of elliptic PDEs, LSFEMs have been used successfully as a starting point for designing multilevel solution techniques with provably optimal behavior [11, 12, 19, 7]. LS methods naturally provide a sharp error estimator [4], which can be used advantageously to design adaptive refinement techniques using composite grids in a multilevel context [25].

Discontinuous FEMs for hyperbolic PDEs, in particular, DG methods [21], have enjoyed substantial interest in recent years [14, 17]. They have proved to be effective and versatile high-order methods for nonlinear hyperbolic systems with natural conservation properties and good monotonicity properties near discontinuities due to upwinding. They can handle nonmatching grids and nonuniform polynomial approximations, orthogonal bases can be chosen that lead to diagonal mass matrices, and they are easily parallelized by using block-type preconditioners [14, 17].

The contributions of the present paper are threefold. First, we establish finite element convergence of the continuous and discontinuous LSFEM formulations proposed in this paper. We start out by presenting a trace theorem that precisely identifies the space of admissible boundary data. Our continuous LSFEM is a modification of the LSFEM studied by Bochev and Choi [6] for a problem similar to (1.1)–(1.2), in which (1.1) is replaced by

$$\mathbf{b} \cdot \nabla p + c\, p = f \quad \text{in } \Omega. \tag{1.4}$$

Their convergence proof for this modified problem does not carry over to our LSFEM formulation for (1.1)–(1.2). Our discontinuous LSFEM (DLSFEM) is a slight modification of the method proposed by Houston, Jensen, and Süli in [17], which does not provide a rigorous finite element convergence proof for this method.

Second, we study the order of convergence of our LSFEM and DLSFEM for discontinuous flow solutions in the general case that the discontinuity is not aligned with the computational mesh. For extensive studies of solution quality and convergence orders for continuous flows, we refer the reader to [5, 6, 17]. Bochev and Choi [5] show in numerical LSFEM experiments that no substantial spurious oscillations arise near discontinuities in the solution. This finding is confirmed in Houston, Jensen, and Süli [17] for DLSFEMs. Both papers show that for continuous flows the accuracy of (D)LSFEMs is comparable to (D)G and (D)SUPG results (especially for higher-order elements and fine grids), while for discontinuous solutions the smearing is substantially larger in the (D)LSFEM results.

In [5, 17], the order of convergence for discontinuous flow solutions is not investigated. In the present paper we study numerical convergence of discontinuous flow solutions for elements of increasing polynomial degree on triangular and quadrilateral meshes. Our numerical study of discontinuous flow simulation with LSFEMs and DLSFEMs yields interesting results. The smearing of the discontinuity improves, while the overshoots and oscillations remain contained as we increase the order of the polynomial degree of the finite elements. We find an increase in the convergence rate as the polynomial degree increases. We observe similar behavior in the L^2 norm and functional norm for LSFEMs and DLSFEMs and for different scalar flow fields.

Third, we study the performance of a standard AMG method [27], which is known to be optimal for large classes of elliptic PDEs. We apply AMG to a conforming LSFEM discretization of the hyperbolic PDE, and we discuss strategies that may overcome some difficulties encountered. The matrices resulting from (D)LSFEMs are SPD, which often is advantageous for the convergence of iterative methods. In particular, the Ruge–Stüben AMG algorithm [27] we use relies on interpolation and

coarsening heuristics that assume SPD matrices. In this paper we treat simple model problems for which there is a time-like direction that can also be exploited by explicit marching schemes. For optimal global implicit solvers, the number of operations per grid point is bounded, which means that even for this kind of time-like problem they may be able to compete with explicit marching methods, for which the number of operations per grid point is bounded as well. This may especially be true when adaptive refinement and derefinement is taken into account. Moreover, explicit marching schemes are limited by time step constraints, which can be severe on adaptively refined grids. Also, efficient parallel implementations have been developed for AMG solvers [16]. Global implicit solves may be especially competitive for the simulation of problems for which there is no preferred marching direction, e.g., steady flows with rotation or flows of mixed elliptic-hyperbolic type. In this paper we make an important first step by investigating whether optimal AMG solvers can be constructed for simple time-like problems. Optimal global solvers for flows without preferred directions will be treated in a forthcoming work.

This paper is organized as follows. In the next section, we examine the space of admissible boundary data (g in (1.2)) and establish a trace theorem and Poincaré inequality. This leads, in section 3, to the formulation of a minimization principle of a LS functional with boundary term, from which a weak form is derived. Coercivity and continuity are proved and a priori estimates are obtained. Well-posedness is also proved for a slightly modified functional that is suitable for computations. In section 4, we describe conforming FEMs that are obtained when the LS functional is minimized over finite dimensional subspaces and error bounds for discontinuous solutions are discussed. In section 5, a DLSFEM is obtained by minimizing a modified functional that incorporates jump terms over a discontinuous finite dimensional space. Section 6 presents a numerical study of the convergence behavior of LSFEMs and DLSFEMs for discontinuous solutions and for elements of increasing polynomial degree on triangular and quadrilateral meshes. The sharpness and monotonicity of the approximate solution in the neighborhood of discontinuities is investigated. In section 7, we study the performance of a standard AMG method [27], which is known to be optimal for large classes of elliptic PDEs. We apply AMG to a conforming LSFEM discretization of the hyperbolic PDE. Conclusions are formulated in section 8.

2. Admissible boundary data. In this section, we examine the space of admissible boundary data for (1.1)–(1.2) and formulate a Poincaré inequality and a trace theorem.

Given Ω in (1.1)–(1.2) $\subset \mathbb{R}^d$, let $\mathbf{b}(\mathbf{x}) = (b_1(\mathbf{x}), \ldots, b_d(\mathbf{x}))$ be a vector field on Ω. We make the following assumptions on $\mathbf{b}$: for any $\hat{\mathbf{x}} \in \Gamma_I$, let $\mathbf{x}(r) = (x_1(r), \ldots, x_d(r))$ be a streamline of $\mathbf{b}$, that is, the solution of

$$\frac{dx_i(r)}{dr} = b_i(\mathbf{x}(r)), \quad i = 1, \ldots, d, \tag{2.1}$$

with initial condition $\mathbf{x}(r_0) = \hat{\mathbf{x}}$. In this paper, we limit the discussion to the case where $d = 2$, although extensions to higher dimensions can be established [24]. Let $\beta = |\mathbf{b}|$, and assume there exist constants β_0 and β_1 such that $0 < \beta_0 \leq \beta \leq \beta_1 < \infty$ on Ω. We assume that there exists a transformation to a coordinate system (r, s) such that the streamlines are lined up with the r coordinate direction and the Jacobian, J, of the transformation is bounded. This implies that no two streamlines intersect and that Ω is the collection of all such streamlines. Furthermore, we assume that every streamline connects Γ_I and Γ_O with a finite length $\ell(\hat{\mathbf{x}})$, where $\hat{\mathbf{x}} \in \Gamma_I$. We

require partition $\mathcal{T}^h$ of Ω to be an admissible, quasi-uniform tessellation (see [8, 9]). We assume the same for $\hat{\mathcal{T}}^h$ of $\hat{\Omega}$, the image of $\mathcal{T}^h$ under the transformation. For our numerical tests, we use uniform partitions of triangles and quadrilaterals.

We define the boundary norm

$$\|g\|_{B_\ell}^2 := \int_{\Gamma_I} \ell(\mathbf{x}(\sigma))|\hat{\mathbf{b}}\cdot\mathbf{n}|g^2 d\sigma, \tag{2.2}$$

where $\hat{\mathbf{b}}$ is the unit vector in the direction $\mathbf{b}$ and $\ell(\mathbf{x})$ is the length of the streamline of $\mathbf{b}$ connecting Γ_I to the outflow boundary Γ_O. Define the space B_ℓ to be the closure of $C^\infty(\Gamma_I)$ in the B_ℓ-norm (2.2). Assuming $f \in L^2(\Omega)$ in (1.1) and using standard notation for L^2 norms, we define the natural norm (often called the *graph* norm) as

$$\|p\|_{V_\ell}^2 := \|p\|_{0,\Omega}^2 + \|\mathbf{b}\cdot\nabla p\|_{0,\Omega}^2 \tag{2.3}$$

and the solution space as

$$V_\ell := \{p \in L^2(\Omega) \ : \ \|p\|_{V_\ell} < \infty\}. \tag{2.4}$$

Remark 2.1. Depending on $\mathbf{b}$ and Ω, B_ℓ can be larger than $L^2(\Gamma_I)$.

LEMMA 2.2 (trace inequality). *If $p \in V_\ell$ and $p = g$ on Γ_I, then there exists a constant C, depending on β_0 and the transformation Jacobian J, such that*

$$\|g\|_{B_\ell}^2 \le C\left(\|p\|_{0,\Omega}^2 + \|\mathbf{b}\cdot\nabla p\|_{0,\Omega}^2\right). \tag{2.5}$$

Proof. We first prove (2.5), assuming $\mathbf{b}$ is constant. Let $\bar{\mathbf{b}} = \frac{1}{|\mathbf{b}|}\mathbf{b}$, the unit vector in the direction of $\mathbf{b}$. For every $\hat{\mathbf{x}} \in \Gamma_I$, let

$$\ell(\hat{\mathbf{x}}) = |s_1(\hat{\mathbf{x}})|, \tag{2.6}$$

where $(0, s_1)$ is the largest interval for which $\hat{\mathbf{x}} + s\hat{\mathbf{b}} \in \Omega$ for all $s \in (0, s_1)$. Here, $\hat{\mathbf{b}} = \bar{\mathbf{b}}(\hat{\mathbf{x}})$ generates the unique streamline intersecting the point $\hat{\mathbf{x}}$. Let $d\sigma$ be the differential arc length along Γ_I and $\mathbf{n}$ the outward unit normal on Γ_I. Then, for any $p \in V_\ell$, we have

$$\iint_\Omega p(\mathbf{x})\, dA = \int_{\Gamma_I}\int_0^{s_1(\hat{\mathbf{x}})} p(\hat{\mathbf{x}} + s\hat{\mathbf{b}})\, ds\, |\bar{\mathbf{b}}\cdot\mathbf{n}|\, d\sigma. \tag{2.7}$$

For any $s \in [0, s_1(\hat{\mathbf{x}})]$, we have

$$p^2(\hat{\mathbf{x}} + s\hat{\mathbf{b}}) = p^2(\hat{\mathbf{x}}) + \int_0^s \bar{\mathbf{b}}\cdot\nabla p^2(\hat{\mathbf{x}} + t\hat{\mathbf{b}})\, dt, \tag{2.8}$$

so

$$p^2(\hat{\mathbf{x}}) \le p^2(\hat{\mathbf{x}} + s\hat{\mathbf{b}}) + \int_0^{s_1(\hat{\mathbf{x}})} \left|\bar{\mathbf{b}}\cdot\nabla p^2(\hat{\mathbf{x}} + t\hat{\mathbf{b}})\right| dt. \tag{2.9}$$

Integrating over $(0, s_1(\hat{\mathbf{x}}))$ with length element dt and using the relation $\ell(\hat{\mathbf{x}}) = \int_0^{s_1(\hat{\mathbf{x}})} dt$, we thus obtain

$$\ell(\hat{\mathbf{x}})p^2(\hat{\mathbf{x}}) \le \int_0^{s_1(\hat{\mathbf{x}})} p^2(\hat{\mathbf{x}} + t\hat{\mathbf{b}})\, dt + \ell(\hat{\mathbf{x}})\int_0^{s_1(\hat{\mathbf{x}})} \left|\bar{\mathbf{b}}\cdot\nabla p^2(\hat{\mathbf{x}} + t\hat{\mathbf{b}})\right| dt. \tag{2.10}$$

Integrating along Γ_I with length element $|\bar{\mathbf{b}} \cdot \mathbf{n}| d\sigma$ yields

$$\begin{aligned}\int_{\Gamma_I} \ell(\hat{\mathbf{x}}) p^2(\hat{\mathbf{x}}) |\bar{\mathbf{b}} \cdot \mathbf{n}| d\sigma \leq & \int_{\Gamma_I} \int_0^{s_1(\hat{\mathbf{x}})} p^2(\hat{\mathbf{x}} + t\hat{\mathbf{b}})\, dt\, |\bar{\mathbf{b}} \cdot \mathbf{n}|\, d\sigma \\ & + \int_{\Gamma_I} \ell(\hat{\mathbf{x}}) \int_0^{s_1(\hat{\mathbf{x}})} \left| \bar{\mathbf{b}} \cdot \nabla p^2(\hat{\mathbf{x}} + t\hat{\mathbf{b}}) \right| dt\, |\bar{\mathbf{b}} \cdot \mathbf{n}|\, d\sigma. \end{aligned} \tag{2.11}$$

Let $D = \mathrm{diam}(\Omega)$. Applying the Cauchy–Schwarz and ε inequalities, we thus have

$$\begin{aligned} \|p\|_{B_\ell}^2 &\leq \|p\|_{0,\Omega}^2 + 2D \|(\bar{\mathbf{b}} \cdot \nabla) p\|_{0,\Omega} \|p\|_{0,\Omega} \\ &\leq \|p\|_{0,\Omega}^2 + D^2 \|p\|_{0,\Omega}^2 + \|(\bar{\mathbf{b}} \cdot \nabla) p\|_{0,\Omega}^2 \\ &\leq C(\|p\|_{0,\Omega}^2 + \|(\mathbf{b} \cdot \nabla) p\|_{0,\Omega}^2). \end{aligned} \tag{2.12}$$

For the general case of variable $\mathbf{b}(\mathbf{x})$, the bound (2.5) follows using the assumed transformation with bounded Jacobian and the fact that $p = g$ on the inflow boundary Γ_I. ∎

Remark 2.3. The constants C which appear in Lemma 2.2 and throughout the rest of the paper are generic and may change value with each occurence but depend only on β_0, Γ_I, and Ω.

LEMMA 2.4 (Poincaré inequality). *Let $D = \mathrm{diam}(\Omega)$. There exists a constant C, depending on β_0 and the transformation Jacobian J, such that*

$$\|p\|_{0,\Omega}^2 \leq C(\|p\|_{B_\ell}^2 + D^2 \|\mathbf{b} \cdot \nabla p\|_{0,\Omega}^2). \tag{2.13}$$

Proof. As in the preceding proof, we derive this Poincaré inequality for constant $\mathbf{b}$ and rely on the transformation with bounded Jacobian to achieve the general result. Let $\bar{\mathbf{b}} = \frac{1}{|\mathbf{b}|}\mathbf{b}$, and let 0 and $s_1(\mathbf{x})$ be as in the proof of Lemma 2.2. For every $\hat{\mathbf{x}} \in \Gamma_I$, let $\ell(\hat{\mathbf{x}}) = |s_1(\hat{\mathbf{x}})|$. Also, let $\hat{\mathbf{b}} = \bar{\mathbf{b}}(\hat{\mathbf{x}})$ generate the unique streamline intersecting the point $\hat{\mathbf{x}}$. Notice that for $s \in [0, s_1(\mathbf{x})]$, we have

$$p(\hat{\mathbf{x}} + s\hat{\mathbf{b}}) = p(\hat{\mathbf{x}}) + \int_0^s \bar{\mathbf{b}} \cdot \nabla p(\hat{\mathbf{x}} + t\hat{\mathbf{b}})\, dt. \tag{2.14}$$

Squaring both sides and using the ε and Jensen inequalities yields

$$\begin{aligned} |p(\hat{\mathbf{x}} + s\hat{\mathbf{b}})|^2 &\leq 2 \left(|p(\hat{\mathbf{x}})|^2 + \left(\int_0^{s_1(\hat{\mathbf{x}})} \left| \bar{\mathbf{b}} \cdot \nabla p(\hat{\mathbf{x}} + t\hat{\mathbf{b}}) \right| dt \right)^2 \right) \\ &\leq 2 \left(|p(\hat{\mathbf{x}})|^2 + \ell(\hat{\mathbf{x}}) \int_0^{s_1(\hat{\mathbf{x}})} \left| \bar{\mathbf{b}} \cdot \nabla p(\hat{\mathbf{x}} + t\hat{\mathbf{b}}) \right|^2 dt \right). \end{aligned} \tag{2.15}$$

Integrating over $(0, s_1(\hat{\mathbf{x}}))$ with dt and using the relation $\ell(\hat{\mathbf{x}}) = \int_0^{s_1(\hat{\mathbf{x}})} dt$ we thus obtain

$$\int_0^{s_1(\hat{\mathbf{x}})} \left| p(\hat{\mathbf{x}} + t\hat{\mathbf{b}}) \right|^2 dt \leq 2 \left(\ell(\hat{\mathbf{x}}) |p(\hat{\mathbf{x}})|^2 + \ell(\hat{\mathbf{x}})^2 \int_0^{s_1(\hat{\mathbf{x}})} \left| \bar{\mathbf{b}} \cdot \nabla p(\hat{\mathbf{x}} + t\hat{\mathbf{b}}) \right|^2 dt \right). \tag{2.16}$$

Integrating along Γ_I with $|\bar{\mathbf{b}} \cdot \mathbf{n}| d\sigma$ and using (2.7) then yields

$$\begin{aligned} \|p\|_{0,\Omega}^2 &\leq 2 \left(\int_{\Gamma_I} \ell(\hat{\mathbf{x}}) p^2(\hat{\mathbf{x}}) |\bar{\mathbf{b}} \cdot \mathbf{n}| d\sigma + D^2 \|\bar{\mathbf{b}} \cdot \nabla p\|_{0,\Omega}^2 \right) \\ &\leq C \left(\|p\|_{B_\ell}^2 + D^2 \|\mathbf{b} \cdot \nabla p\|_{0,\Omega}^2 \right). \end{aligned} \tag{2.17}$$

∎

The following trace theorem establishes B_ℓ as the space of admissible functions for inflow boundary conditions when the right-hand side, f, in (1.1) is in $L^2(\Omega)$.

THEOREM 2.5 (trace theorem). *For $p \in V_\ell$ let $\gamma(p)$ represent the trace of p on Γ_I. Then the map $\gamma : V_\ell \to B_\ell$ is a bounded surjection.*

Proof. For any $g \in B_\ell$, we can construct a flat function p such that $p = g$ on Γ_I and $\mathbf{b} \cdot \nabla p = 0$ in Ω. From the Poincaré inequality (2.13), it follows that $p \in V_\ell$. Together with the trace inequality, this yields the trace theorem. ∎

Remark 2.6. Our trace theorem is similar to the theorem proved in [24] for the more general case of the neutron transport equation in phase space. A different characterization of the trace space is given in [18] for the general class of Friedrichs systems, of which (1.1)–(1.2) is a special case. The trace operator defined in [18] is not surjective. In this sense, in contrast to Theorem 2.5, the trace space identified in [18] does not provide a sharp trace theorem.

3. LS weak form. In this section, we formulate a LS minimization principle, derive the weak form of the minimization, and prove existence of a unique $p \in V_\ell$ solving the weak problem. We use the tools developed in the previous section and coercivity and continuity with respect to the natural norm (2.3) to arrive at these results.

We define the LS functional

$$\mathcal{G}_\ell(p;\, f, g) := \|\mathbf{b} \cdot \nabla p - f\|^2_{0,\Omega} + \|p - g\|^2_{B_\ell}. \tag{3.1}$$

First we note that if p satisfies (1.1)–(1.2), then

$$p = \arg\min_{p \in V_\ell} \mathcal{G}_\ell(p;\, f, g).$$

The bilinear form associated with $\mathcal{G}_\ell$ (3.1) is

$$\mathcal{F}_\ell(p, q) := \langle \mathbf{b} \cdot \nabla p,\, \mathbf{b} \cdot \nabla q \rangle_{0,\Omega} + \langle p,\, q \rangle_{B_\ell},$$

with standard notation for scalar products associated with norms. The weak form of the minimization is as follows.

PROBLEM 3.1. *Find $p \in V_\ell$ s.t.*

$$\mathcal{F}_\ell(p, q) = F(q) \qquad \forall\, q \in V_\ell, \tag{3.2}$$

where

$$F(q) = \langle f,\, \mathbf{b} \cdot \nabla q \rangle_{0,\Omega} + \langle g,\, q \rangle_{B_\ell}.$$

Note that $F(q) \in V_\ell'$, the dual space of V_ℓ.

The following establishes coercivity and continuity in the V_ℓ norm of the bilinear form, $\mathcal{F}_\ell(\cdot, \cdot)$, defined by (3.2). With these properties the bilinear form, $\mathcal{F}_\ell(\cdot, \cdot)$, is frequently referred to as *V_ℓ-elliptic* [8].

THEOREM 3.2 (coercivity and continuity, existence and uniqueness). *There exist constants c_0 and c_1 s.t. for every $p, q \in V_\ell$*

$$c_0 \|p\|^2_{V_\ell} \le \mathcal{F}_\ell(p, p), \tag{3.3}$$

$$\mathcal{F}_\ell(p, q) \le c_1 \|p\|_{V_\ell} \|q\|_{V_\ell}. \tag{3.4}$$

Furthermore, for every $f \in L^2(\Omega)$, $g \in B_\ell$, there exists a unique $p \in V_\ell$ solving the weak problem (3.2)*. Moreover, p also satisfies* (1.1)–(1.2)*.*

Proof. Using the definition of $\|p\|_{V_\ell}$ from (2.3) and Poincaré inequality (2.13) yields

$$\begin{aligned} \|p\|_{V_\ell}^2 &\le C(\|p\|_{B_\ell}^2 + D^2\|\mathbf{b}\cdot\nabla p\|_{0,\Omega}^2) + \|\mathbf{b}\cdot\nabla p\|_{0,\Omega}^2 \\ &\le C(\|p\|_{B_\ell}^2 + \|\mathbf{b}\cdot\nabla p\|_{0,\Omega}^2) \\ &= C\mathcal{F}_\ell(p,p), \end{aligned} \tag{3.5}$$

which yields (3.3). Similarly, applying the Cauchy–Schwarz inequality followed by trace inequality (2.5) and Cauchy–Schwarz again, we have

(3.6)

$$\begin{aligned} \mathcal{F}_\ell(p,q) &\le \|\mathbf{b}\cdot\nabla p\|_{0,\Omega}\|\mathbf{b}\cdot\nabla q\|_{0,\Omega} + \|p\|_{B_\ell}\|q\|_{B_\ell} \\ &\le \|\mathbf{b}\cdot\nabla p\|_{0,\Omega}\|\mathbf{b}\cdot\nabla q\|_{0,\Omega} + C\big(\|p\|_{0,\Omega}^2 + \|\mathbf{b}\cdot\nabla p\|_{0,\Omega}^2\big)^{\frac{1}{2}} C\big(\|q\|_{0,\Omega}^2 + \|\mathbf{b}\cdot\nabla q\|_{0,\Omega}^2\big)^{\frac{1}{2}} \\ &\le C\|p\|_{V_\ell}\|q\|_{V_\ell}, \end{aligned}$$

which confirms (3.4).

The trace theorem and the Cauchy–Schwarz inequality imply that, for every $f \in L^2(\Omega)$ and $g \in B_\ell$,

$$F(q) := \langle f,\, \mathbf{b}\cdot\nabla q\rangle_{0,\Omega} + \langle g,\, q\rangle_{B_\ell} \tag{3.7}$$

is a bounded linear functional on V_ℓ. Thus, we can embed the pair $(f,g) \in L^2(\Omega)\times B_\ell$ into V_ℓ', the dual space of V_ℓ.

By the Lax–Milgram theorem [8], for all $(f,g) \in L^2(\Omega)\times B_\ell$, there exists a unique $p \in V_\ell$ that satisfies the weak problem (3.2). We now show that p also solves the strong problem (1.1)–(1.2). It suffices to show that the embedding of $L^2(\Omega)\times B_\ell$ into V_ℓ' is injective.

To do this, pick $(f,g) \in L^2(\Omega)\times B_\ell$ and suppose

$$F(q) = \langle f,\, \mathbf{b}\cdot\nabla q\rangle_{0,\Omega} + \langle g,\, q\rangle_{B_\ell} = 0 \tag{3.8}$$

for every $q \in V_\ell$. Thus, if $f = 0$ and $g = 0$, the embedding is injective.

We first show that for $(f,g) \in L^2(\Omega)\times B_\ell$, there exists $p_s \in V_\ell$ such that

$$\mathcal{L}p_s = (f,g), \tag{3.9}$$

where $\mathcal{L} : V_\ell \to L^2(\Omega)\times B_\ell$ is defined by

$$\mathbf{b}\cdot\nabla p_s = f \quad \text{in } \Omega, \tag{3.10a}$$

$$p_s = g \quad \text{on } \Gamma_I. \tag{3.10b}$$

That is, we must show that $\mathcal{L}$ is surjective. Construct p_s as follows. Let p_1 be a flat function such that

$$\mathbf{b}\cdot\nabla p_1 = 0 \quad \text{in } \Omega, \tag{3.11a}$$

$$p_1 = g \quad \text{on } \Gamma_I. \tag{3.11b}$$

For $\mathbf{x} \in \Omega$, let $\hat{\mathbf{x}}$ be the point on Γ_I with the same streamline as $\mathbf{x}$. Let $\beta(\mathbf{x}) = |\mathbf{b}(\hat{\mathbf{x}})|$ and $\bar{\mathbf{b}} = \frac{\mathbf{b}(\mathbf{x})}{\beta(\mathbf{x})}$. Let p_2 be given by

$$p_2(\mathbf{x}) = \int_{\hat{\mathbf{x}}}^{\mathbf{x}} \frac{f(\hat{\mathbf{x}} + s\bar{\mathbf{b}}(\hat{\mathbf{x}}))}{\beta(\hat{\mathbf{x}} + s\bar{\mathbf{b}}(\hat{\mathbf{x}}))}\, ds. \tag{3.12}$$

Then, p_2 satisfies

$$\mathbf{b}\cdot\nabla p_2 = f \quad \text{in } \Omega, \tag{3.13a}$$

$$p_2 = 0 \quad \text{on } \Gamma_I. \tag{3.13b}$$

Writing $p_s = p_1 + p_2$, we see that p_s satisfies (3.10a)–(3.10b). Thus, $\mathcal{L}$ is a surjection.

Now since $F(q) = 0$ and $p_s \in V_\ell$ satisfies (3.10a)–(3.10b), we have

$$\langle f, f\rangle_{0,\Omega} + \langle g,\, g\rangle_{B_\ell} = 0. \tag{3.14}$$

Thus, $f = 0$ and $g = 0$. It follows that the embedding is injective. This completes the proof. ∎

The following a priori estimate is a direct consequence of Theorem 3.2. These bounds are often referred to as stability estimates.

COROLLARY 3.3 (a priori estimate). *There exist constants c_3 and c_4 such that if p satisfies* (3.2)*, then*

$$c_3\|p\|_{V_\ell} \le (\|f\|_{0,\Omega} + \|g\|_{B_\ell}) \le c_4\|p\|_{V_\ell}. \tag{3.15}$$

Proof. The proof follows directly from Theorem 3.2. ∎

For certain problems, $\ell(\mathbf{x})$ in (2.2) may not be easily computed, making the LS formulation intractable. To avoid this difficulty we modify the boundary norm in the functional to be

$$\|g\|_B^2 := \int_{\Gamma_I} |\hat{\mathbf{b}}\cdot\mathbf{n}| g^2 ds, \tag{3.16}$$

where $\hat{\mathbf{b}}$ is the unit normal in the direction of $\mathbf{b}$. Let $B = \{g : \|g\|_B < \infty\}$, and notice that $B_\ell \cap L^\infty(\Gamma_I) \subseteq B$. If Ω is such that Γ_I and Γ_O remain a bounded distance apart, then this norm is equivalent to the original norm, $\|\cdot\|_{B_\ell}$. If Γ_I and Γ_O touch, then there are functions in B_ℓ that are not in B. For bounded functions, the B_ℓ norm and B norm are equivalent. That is, if we restrict our attention to bounded boundary data, then nothing is lost in modifying the functional.

In general, the trace inequality (2.5) does not hold with B_ℓ replaced by B, but the Poincaré inequality (2.13) does. To retain the inequalities and ellipticity results obtained above, we must include the boundary term in the definition of the norm. Define the norm

$$\|p\|_V^2 := \|p\|_{0,\Omega}^2 + \|\mathbf{b}\cdot\nabla p\|_{0,\Omega}^2 + \|p\|_B^2$$

and the space

$$V := \{p \in L^2(\Omega) \;:\; \|p\|_V < \infty\}.$$

The modified functional is then defined as follows: let $f \in L^2(\Omega)$, let $g \in B$, and define

$$\mathcal{G}(p;\, f, g) := \|\mathbf{b}\cdot\nabla p - f\|_{0,\Omega}^2 + \|p - g\|_B^2. \tag{3.17}$$

If p satisfies (1.1)–(1.2), then

$$p = \arg\min_{p\in V} \mathcal{G}(p;\, f, g).$$

The associated bilinear form is

$$\mathcal{F}(p,q) := \langle \mathbf{b}\cdot\nabla p,\ \mathbf{b}\cdot\nabla q\rangle_{0,\Omega} + \langle p,\ q\rangle_B, \tag{3.18}$$

and the weak form of the minimization is as follows.

PROBLEM 3.4. *Find $p \in V$ s.t.*

$$\mathcal{F}(p,q) = F(q) \qquad \forall\, q \in V, \tag{3.19}$$

where

$$F(q) = \langle f,\ \mathbf{b}\cdot\nabla q\rangle_{0,\Omega} + \langle g,\ q\rangle_B. \tag{3.20}$$

With this change, we obtain existence and uniqueness and an a priori estimate as before.

THEOREM 3.5 (coercivity and continuity, existence and uniqueness). *There exist constants c_0 and c_1 s.t. for every $p, q \in V$*

$$c_0\|p\|_V^2 \le \mathcal{F}(p,p),$$
$$\mathcal{F}(p,q) \le c_1\|p\|_V\|q\|_V.$$

Furthermore, for $f \in L^2$ and $g \in B$, there exists a unique $p \in V$ solving Problem 3.4.

COROLLARY 3.6 (a priori estimates). *There exist constants c_3 and c_4 such that if p satisfies* (3.19)*, then*

$$c_3\|p\|_V \le (\|f\|_{0,\Omega} + \|g\|_B) \le c_4\|p\|_V. \tag{3.21}$$

Remark 3.7. The $\mathcal{G}$ norm, defined as

$$\|p^h\|_{\mathcal{G}}^2 := \mathcal{G}(p^h, 0, 0), \tag{3.22}$$

is a natural and computable a posteriori error estimator. To see this, let $e = p^h - p$, where p solves (1.1)–(1.2). Then

$$\begin{aligned}\|e\|_{\mathcal{G}} &= \mathcal{G}(p^h - p; 0, 0)\\ &= \mathcal{G}(p^h; f, g).\end{aligned}$$

LS methods offer the advantage of a convenient a posteriori error indicator. Sharpness is addressed in [4, 7].

4. Conforming finite elements. In this section we discuss the discrete form of the minimization. We consider an admissible, quasi-uniform tessellation $\mathcal{T}^h$ of Ω (cf. [8]). For a conforming method, we choose the discrete space $V^h \subset V$. For example, in our numerical tests we use uniform partitions of triangles and quadrilaterals and implement piecewise polynomials with continuity imposed across element edges. Let

$$V^h := \mathcal{M}_k^h \cap \mathcal{C}^0(\Omega), \tag{4.1}$$

where

$$\mathcal{M}_k^h := \{p\colon p \in \mathcal{P}_k(\tau)\,\forall\,\tau \in \mathcal{T}^h\}. \tag{4.2}$$

Here, $\mathcal{P}_k(\tau)$ is the space of polynomials of total degree $\le k$ when τ is a triangle and tensor product polynomials of degree $\le k$ in each coordinate direction when τ is a quadrilateral. We now pose the conforming discrete weak form of the minimization.

PROBLEM 4.1. *Find* $p^h \in V^h$ *s.t.*

$$\mathcal{F}(p^h, q^h) = F(q^h) \qquad \forall q^h \in V^h, \tag{4.3}$$

where $\mathcal{F}$ *is defined by* (3.18) *and* F *is defined by* (3.20).

By Ceá's lemma, we have

$$\|p - p^h\|_V \leq \frac{c_0}{c_1} \inf_{\hat{p}^h \in V^h} \|p - \hat{p}^h\|_V,$$

where c_1 and c_0 are the constants from the continuity and coercivity bounds.

In this paper we are interested in discontinuous solutions, p. Suppose g is discontinuous but piecewise smooth. That is, $g \in H^{\frac{1}{2}-\varepsilon}(\Gamma_I)$. Then, for smooth f, p has the same smoothness, $p \in H^{\frac{1}{2}-\varepsilon}(\Omega)$. In this case it can be shown that, for grid-aligned flow,

$$\|p - p^h\|_V \leq Ch^{\frac{1}{2}-\varepsilon}\|p\|_{\frac{1}{2}-\varepsilon},$$

where C is some grid-independent constant. The exact bound for the non-grid-aligned case remains an open question. Still, the theoretical limit for the grid-aligned case and other results offer some insight. Scott and Zhang describe in [28] an interpolation $\tilde{I}^h$ such that $\|p - \tilde{I}^h p\|_{0,\Omega} \leq Ch^{\frac{1}{2}-\varepsilon}\|p\|_{\frac{1}{2}-\varepsilon}$. If we assume that $\frac{1}{2}$ is the optimal L^2-rate of convergence for interpolation, then we expect that the L^2-rate of convergence for the FEM will be no better than $\frac{1}{2}$. Note that the Poincaré inequality (2.13) yields $\|p - p^h\|_{0,\Omega} \leq C\|p - p^h\|_V$. Thus, the V norm rate of convergence cannot be faster than the L^2-norm rate. In section 6, we discuss our numerical findings regarding error estimates and present results consistent with the error bounds proposed. We find that, as we increase the order of the elements, the convergence rate increases and is bounded by $\frac{1}{2}$ in both the L^2 norm and $\mathcal{G}$ norm. For an extensive analysis of error bounds and convergence rates for smooth solutions see [5, 6].

5. Nonconforming finite elements. In this section we describe the use of discontinuous elements motivated by the case when the flow is grid-aligned. Consider an example when the characteristics follow the grid and the boundary data is prescribed such that the discontinuity in the solution follows the element edges aligned with the characteristics. In (1.1)–(1.2) let $f = 0$ and prescribe piecewise constant boundary data with discontinuities only at nodes. If we use the discontinuous space $\mathcal{M}_k^h$ defined by (4.2), the solution to (1.1)–(1.2) is in this space. However, the grids we consider are generally not aligned with the flow field $\mathbf{b}(\mathbf{x})$, and boundary data is often more general than in this special case. If attention is given to the behavior of the jumps with respect to the grid, a well-posed formulation of the problem in a discontinuous LS setting is attainable. To this end, let $\mathcal{T}^h = \bigcup_j \tau_j$ be a tessellation of Ω, and let $S^h := \mathcal{M}_k^h$ be defined as in (4.2). Let $\Gamma_{i,j} := \tau_i \cap \tau_j$ denote the edge common to elements τ_i and τ_j. Since $S^h \not\subset V$, we call S^h a nonconforming space [8].

For $p^h \in S^h + V$ define the element edge functional as

$$\|p^h\|_{E^h}^2 := \sum_{i,j} \omega_{i,j} \int_{\Gamma_{i,j}} |\mathbf{b} \cdot \mathbf{n}_\tau| [\![p^h]\!]^2 ds. \tag{5.1}$$

Here, $\mathbf{n}_\tau$ is the outward unit normal to edge $\Gamma_{i,j}$, $\omega_{i,j}$ is a weight to be determined, and $[\![p^h]\!]$ is the jump in p^h across $\Gamma_{i,j}$. We use the term (5.1) in the LS functional to

make a distinction between element edges that are closely aligned with the flow and edges that are not by tying together neighboring elements. This behavior is consistent with the regularity of the solution. A solution p of (1.1)–(1.2) would be smooth in the direction of the flow while perpendicular to the flow p is only L^2-regular. For further motivation, consider a non-grid-aligned flow with a typical discontinuity (see Figure 1). When element edges are nearly aligned with the discontinuity (location A), the term $|\mathbf{b}\cdot\mathbf{n}|$ is small in the term (5.1), allowing a larger jump between the neighboring elements. However, when an element edge is nearly perpendicular to the flow (location B), $|\mathbf{b}\cdot\mathbf{n}|$ is large. This enforces a stronger connection between the elements resulting in a smaller jump.

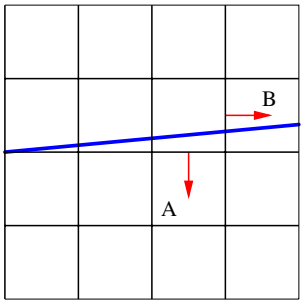

FIG. 1. *Example of a non-grid-aligned flow and outward normals A and B.*

We can now define a nonconforming LS functional similar to (3.17) except for the use of broken norms and inclusion of the edge functional (5.1). With $f \in L^2(\Omega)$, $g \in B$, $p \in S^h + V$, define the functional as

$$\mathcal{G}^h(p; f, g) := \sum_j \|\mathbf{b}\cdot\nabla p - f\|^2_{0,\tau_j} + \|p\|^2_{E^h} + \|p - g\|^2_B. \tag{5.2}$$

Define the $\mathcal{G}^h$ norm as

$$\|p^h\|^2_{\mathcal{G}^h} := \mathcal{G}^h(p^h, 0, 0). \tag{5.3}$$

If $p \in V$, then $\mathcal{G}^h(p; f, g) = \mathcal{G}(p; f, g)$.

Let $e = p^h - p$, where $p^h \in S^h$ and p satisfies (1.1)–(1.2). Notice that

$$\begin{aligned}\|e\|_{\mathcal{G}^h} &= \mathcal{G}^h(p^h - p; 0, 0)\\ &= \mathcal{G}^h(p^h; f, g).\end{aligned}$$

Thus, $\mathcal{G}^h$ is a natural a posteriori error estimator. The sharpness of LS error estimators is addressed in [4, 7].

We can now describe the discrete variational problem for our discontinuous elements.

PROBLEM 5.1. *Find $p^h \in S^h$ s.t.*

$$\mathcal{F}(p^h, q^h) = F(q^h) \qquad \forall\, q^h \in S^h,$$

where

$$\begin{aligned}\mathcal{F}(p^h, q^h) &:= \sum_{\tau_i} \langle \mathbf{b}\cdot\nabla p^h, \mathbf{b}\cdot\nabla q^h\rangle_{0,\tau_i} + \langle p^h,\, q^h\rangle_{E^h} + \langle p^h,\, q^h\rangle_B,\\ F(q^h) &:= \sum_{\tau_i} \langle f, \mathbf{b}\cdot\nabla q^h\rangle_{0,\tau_i} + \langle g,\, q^h\rangle_B.\end{aligned}$$

In the following lemma we find that a uniform Poincaré inequality is satisfied for weights stronger than $\omega = c\frac{1}{h}$, where c is a grid-independent constant. We also show, by example, that weights weaker than $\omega = c\frac{1}{h}$—e.g., $\omega = 1$ or h—result in a violation of the uniform Poincaré inequality. Thus, enforcing the connection between neighboring elements too weakly not only decreases the stability of the solution but also results in losing a uniform bound on the error in the L^2 norm.

LEMMA 5.2 (uniform Poincaré inequality). *There exists a constant C, independent of h, such that for $p^h \in S^h + V$ and $\omega \geq c\frac{1}{h}$, where c is a grid-independent constant,*

$$\|p^h\|_{0,\Omega} \leq C\|p^h\|_{\mathcal{G}^h}. \tag{5.4}$$

Furthermore, the above does not hold for $\omega < c\frac{1}{h}$.

Proof. Similarly to the proof of Lemma 2.4, we derive the uniform Poincaré inequality for constant $\mathbf{b}$ and rely on the transformation with bounded Jacobian to achieve the general result. As before, let $\bar{\mathbf{b}} = \frac{1}{|\mathbf{b}|}\mathbf{b}$. Let $\hat{\mathbf{x}} \in \Gamma_I$, and let s_k be parameters in $(0, s_m(\hat{\mathbf{x}}))$ such that $\hat{\mathbf{x}}_k = \hat{\mathbf{x}} + s_k\bar{\mathbf{b}}(\hat{\mathbf{x}})$ lies on an element edge, where (s_0, s_m) now plays the role of $(0, s_1)$ in our previous proofs. Since the flow field $\mathbf{b}$ is constant, we have $m(\hat{\mathbf{x}}) = \mathcal{O}(\sqrt{N})$, where N is the number of elements in $\mathcal{T}^h$, the tessellation of Ω, and $m(\hat{\mathbf{x}})$ is the number of element edges encountered by the characteristics generated by $\hat{\mathbf{b}} = \bar{\mathbf{b}}(\hat{\mathbf{x}})$ emanating from $\hat{\mathbf{x}} \in \Gamma_I$. For $0 \leq k < m$, we assume

$$|s_{k+1}(\hat{\mathbf{x}}) - s_k(\hat{\mathbf{x}})| < \tilde{h} \tag{5.5}$$

for all $\hat{\mathbf{x}} \in \Gamma_I$, where

$$\tilde{h} = \max_j\{\text{diam } \tau_j : \tau_j \in \mathcal{T}^h\}. \tag{5.6}$$

Furthermore, assume $\tilde{h} = \mathcal{O}(\frac{1}{\sqrt{N}})$, and let

$$\ell(\hat{\mathbf{x}}) = \sum_{k=1}^{m} |s_k(\hat{\mathbf{x}}) - s_{k-1}(\hat{\mathbf{x}})|. \tag{5.7}$$

Let $[\![p(x)]\!]$ denote the jump in p at x. Using

$$p(\hat{\mathbf{x}} + s\hat{\mathbf{b}}) = p(\hat{\mathbf{x}}) + \sum_{j=1}^{k}\int_{s_{j-1}}^{s_j} \bar{\mathbf{b}} \cdot \nabla p(\hat{\mathbf{x}} + t\hat{\mathbf{b}})\, dt + [\![p(\hat{\mathbf{x}}_j)]\!] + \int_{s_k}^{s} \bar{\mathbf{b}} \cdot \nabla p(\hat{\mathbf{x}} + t\hat{\mathbf{b}})\, dt, \tag{5.8}$$

taking absolute values, extending the range of integration, and then squaring both sides, we arrive at

$$\left|p(\hat{\mathbf{x}} + s\hat{\mathbf{b}})\right|^2 \leq \left(|p(\hat{\mathbf{x}})| + \sum_{j=1}^{m}\int_{s_{j-1}}^{s_j} \left|\bar{\mathbf{b}} \cdot \nabla p(\hat{\mathbf{x}} + t\hat{\mathbf{b}})\right| dt + \sum_{j=1}^{m-1} [\![p(\hat{\mathbf{x}}_j)]\!]\right)^2. \tag{5.9}$$

Using the inequality

$$\left(\sum_{j=1}^{M} a_j\right)^2 \leq M\sum_{j=1}^{M} a_j^2, \tag{5.10}$$

(5.5), and Jensen's inequality, we obtain

(5.11)
$$\begin{aligned}\left|p(\hat{\mathbf{x}}+s\hat{\mathbf{b}})\right|^2 &\le 3\left\{|p(\hat{\mathbf{x}})|^2+\left(\sum_{j=1}^{m}\int_{s_{j-1}}^{s_j}\left|\bar{\mathbf{b}}\cdot\nabla p(\hat{\mathbf{x}}+t\hat{\mathbf{b}})\right|\,dt\right)^2+\left(\sum_{j=1}^{m-1}[\![p(\hat{\mathbf{x}}_j)]\!]\right)^2\right\}\\ &\le 3\left\{|p(\hat{\mathbf{x}})|^2+m\sum_{j=1}^{m}\tilde{h}\int_{s_{j-1}}^{s_j}\left|\bar{\mathbf{b}}\cdot\nabla p(\hat{\mathbf{x}}+t\hat{\mathbf{b}})\right|^2\,dt+m\sum_{j=1}^{m-1}[\![p(\hat{\mathbf{x}}_j)]\!]^2\right\}.\end{aligned}$$

Using the fact that $m\tilde{h}\le CD$, where $D=\operatorname{diam}(\Omega)$, and integrating over $\int_0^{s_m}dt$, we have

(5.12)
$$\begin{aligned}\sum_{j=1}^{m}\int_{s_{j-1}}^{s_j}\left|p(\hat{\mathbf{x}}+t\hat{\mathbf{b}})\right|^2dt \le 3\Big\{&\ell(\hat{\mathbf{x}})\,|p(\hat{\mathbf{x}})|^2+CD\ell(\hat{\mathbf{x}})\sum_{j=1}^{m}\int_{s_{j-1}}^{s_j}\left|\bar{\mathbf{b}}\cdot\nabla p(\hat{\mathbf{x}}+t\hat{\mathbf{b}})\right|^2dt\\ &+m\ell(\hat{\mathbf{x}})\sum_{j=1}^{m-1}[\![p(\hat{\mathbf{x}}_j)]\!]^2\Big\}.\end{aligned}$$

We now integrate according to $\int_{\Gamma_I}\cdot|\bar{\mathbf{b}}\cdot\mathbf{n}|\,d\sigma$ to get

(5.13)
$$\begin{aligned}&\int_{\Gamma_I}\sum_{j=1}^{m}\int_{s_{j-1}}^{s_j}\left|p(\hat{\mathbf{x}}+t\hat{\mathbf{b}})\right|^2dt\,\left|\bar{\mathbf{b}}\cdot\mathbf{n}\right|\,d\sigma\\ &\quad\le 3\Big\{\int_{\Gamma_I}\ell(\hat{\mathbf{x}})(p(\hat{\mathbf{x}}))^2\left|\bar{\mathbf{b}}\cdot\mathbf{n}\right|\,d\sigma+CD^2\int_{\Gamma_I}\sum_{j=1}^{m}\int_{s_{j-1}}^{s_j}\left|\bar{\mathbf{b}}\cdot\nabla p(\hat{\mathbf{x}}+t\hat{\mathbf{b}})\right|^2dt\,\left|\bar{\mathbf{b}}\cdot\mathbf{n}\right|d\sigma\\ &\qquad+m\int_{\Gamma_I}\ell(\hat{\mathbf{x}})\sum_{j=1}^{m-1}[\![p(\hat{\mathbf{x}}_j)]\!]^2\left|\bar{\mathbf{b}}\cdot\mathbf{n}\right|\,d\sigma\Big\}\\ &\quad\le 3\Big\{\|p\|_B^2+\frac{CD^2}{\beta_0}\sum_j\|\mathbf{b}\cdot\nabla p\|_{0,\tau_j}^2+mD\sum_{i,j}\int_{\Gamma_{i,j}}[\![p]\!]^2\left|\mathbf{b}\cdot\mathbf{n}\right|\,ds\Big\}.\end{aligned}$$

If $\omega\le c\frac{1}{h}=\mathcal{O}(m)$, then

$$\begin{aligned}\|p\|_{0,\Omega}^2 &\le C\left\{\|p\|_B^2+\sum_j\|\mathbf{b}\cdot\nabla p\|_{0,\tau_j}^2+\|p\|_{E^h}^2\right\}\\ &= C\|p\|_{\mathcal{G}^h}.\end{aligned}\tag{5.14}$$

For the general case, bound (5.4) now follows using the assumed transformation with bounded Jacobian.

To show that c is not grid independent for $\omega\le c\frac{1}{h}$, consider the example of a "stair-step function." Let $\Omega=[0,1]\times[0,1]$ and partition $\mathcal{T}^h$ be a uniform tessellation of squares. Let $\mathbf{b}=(1,0)^T$ and $h=\frac{1}{N}$, where N is the number of elements in each coordinate direction. Define $p(x,y)$ on Ω as

$$p(x,y)=jh\quad\text{for } x\in[(j-1)h,\,jh)\,,\ j=1,\dots,N.\tag{5.15}$$

Then

$$\|p\|_{0,\Omega}^2 = \mathcal{O}(1) \tag{5.16}$$

and

$$\|p\|_{\mathcal{G}^h}^2 = \mathcal{O}(\omega \cdot h). \tag{5.17}$$

So, unless $\omega \geq c\frac{1}{h}$, inequality (5.4) is violated for grid-independent c. ∎

Remark 5.3. Once the uniform Poincaré inequality is established, Strang's second lemma [8] can be invoked to prove convergence of DLSFEMs. In the absence of the uniform Poincaré inequality, one cannot guarantee that convergence in the grid-dependent norm implies finite element convergence, as illustrated by the "stair-step" example described in the proof above.

Since $V^h \subset S^h$, we can also conclude for $\hat{p}^h \in V^h$ that

$$\|p - p^h\|_{\mathcal{G}^h} = \inf_{\hat{p}^h \in V^h} \|p - \hat{p}^h\|_{\mathcal{G}^h}.$$

Thus, in the $\mathcal{G}^h$ norm the nonconforming solution is at least as small as the solution from the conforming space. This might lead one to believe that the discretization error in the L^2 norm for the nonconforming solution would be smaller than the L^2 error in the conforming solution. However, our numerical tests show that this is not the case. Using the weight $\omega = \frac{1}{h}$ for non-grid-aligned flow, we show numerically that the convergence rates, for both conforming and nonconforming approximations, appear to be increasing, but to be bounded by $\frac{1}{2}$, in both the L^2 norm and $\mathcal{G}^h$ norm as k, the order of the polynomial, increases.

6. Numerical results. In this section we present numerical results in support of our theoretical error estimates and conjectures of sections 4 and 5, and to demonstrate properties of the LS solution in terms of oscillations and smearing. Convergence rates presented in this section are obtained on sequences of grids ranging from $h = 2^{-4}$ to 2^{-9} in mesh size depending on the order of the polynomial, k.

Consider (1.1)–(1.2), and let $\Omega = [0,1] \times [0,1]$. Let $\mathbf{b}(\mathbf{x}) = (\cos(\theta), \sin(\theta))$, where θ is the angle the flow makes with the first coordinate axis. The inflow boundary defined by (1.3) is $\Gamma_I = (\{0\} \times [0,1]) \cup ([0,1] \times \{0\})$—i.e., the west and south boundaries of the unit square. Let $g(0,y) = 1$ and $g(x,0) = 0$ so that the exact solution is discontinuous with $p = 1$ above the characteristic emanating from the origin and $p = 0$ below the characteristic. For the tessellation $\mathcal{T}^h$ of Ω we choose a uniform partition of quadrilaterals and a uniform partition of triangles.

Tables 1 and 2 show that we achieve consistent convergence rates in the L^2 and $\mathcal{G}^h$ norms both for the quadrilateral and triangular elements. Furthermore, as the order of the polynomials increases, the convergence rates seem to be increasing but to be bounded by $\frac{1}{2}$. Figure 2 shows that for increasing degree k the convergence rate (slope) improves slightly, and higher-order methods exhibit smaller error constants per degree of freedom. This suggests that a combination of h and p refinement (where p is the polynomial order) [17] may work well for the kind of discontinuous hyperbolic flows we consider in this paper.

The nonconforming space S^h discussed in section 5 offers the ability for the approximation p^h to be discontinuous at the element edges, with a possibility of leading to faster convergence rates for the interior term in the functional. This is indeed the

TABLE 1
Convergence rates for $\theta = \frac{\pi}{8}$ using quadrilaterals.

	Conforming (4.1)		Nonconforming (4.2) $\omega = \frac{1}{h}$	
k	L^2 norm	$\mathcal{G}$ norm	L^2 norm	$\mathcal{G}^h$ norm
1	.25	.26	.24	.26
2	.34	.33	.32	.33
3	.36	.37	.36	.34
4	.38	.38	.37	.37

TABLE 2
Convergence rates for $\theta = \frac{\pi}{8}$ using triangles.

	Conforming (4.1)		Nonconforming (4.2) $\omega = \frac{1}{h}$	
k	L^2 norm	$\mathcal{G}$ norm	L^2 norm	$\mathcal{G}^h$ norm
1	.25	.28	.23	.24
2	.33	.32	.33	.33
3	.39	.37	.38	.42

case, as shown in Table 3. Since the uniform Poincaré inequality (5.4) does not hold for weaker values of ω, we should expect the $\mathcal{G}^h$ norm to outperform the L^2 norm for weak ω. Moreover, we find that the convergence rates for each term in the functional become less balanced as ω is chosen away from $\frac{1}{h}$.

It is also interesting to study the effect of varying the weight of the boundary functional, e.g., for the continuous LSFEM. Figure 3 shows the convergence order for the L^2 and $\mathcal{G}^h$ norms as a function of boundary functional weight. Only for a weight equal to 1 are the convergence rates in balance, in accordance with our theoretical results in sections 2 and 3.

The above results were obtained using $\theta = \frac{\pi}{8}$. Table 4 reveals that the convergence rates were generally relatively independent of the angle θ. Table 5 shows that for *very* small angles—e.g., $\theta \leq .05$—we find convergence rates very close to $\frac{1}{2}$ for both the L^2 norm and the $\mathcal{G}^h$ norm using conforming and nonconforming elements. Furthermore, as expected, the convergence rates do not exceed $\frac{1}{2}$, and, for the case of grid-aligned flow, the convergence rates are exactly $\frac{1}{2}$.

Smearing of discontinuities is an important consideration for numerical approximation of hyperbolic PDEs. In the exact solution of the model problem, the discontinuity on the inflow boundary Γ_I is advected to the outflow boundary Γ_O without diffusion. However, in a discrete space over a grid that is not flow aligned, we cannot exactly resolve the discontinuity and the finite element solution displays smearing along the characteristic defining the discontinuity.

It is shown in [5, 17] that the (D)LS solution smears the discontinuity substantially more than the SUPG solution, while the Galerkin solution had the least smearing. However, the Galerkin solution exhibits the most oscillations. The SUPG solution exhibits a small amount of oscillation, while the LS solution has almost no oscillation. Oscillations are an impediment to accurate local adaptive refinements, as they obscure where the adaptivity is most effective. Figure 4 confirms these results for our LS methods and also indicates that the smearing decreases for higher-order elements. Nearly identical plots were obtained using nonconforming elements and have been omitted for brevity.

Next we evaluate the oscillations arising in the discrete solution and observe the magnitude of the overshoots. Higher-order elements produce undesirable overshoots

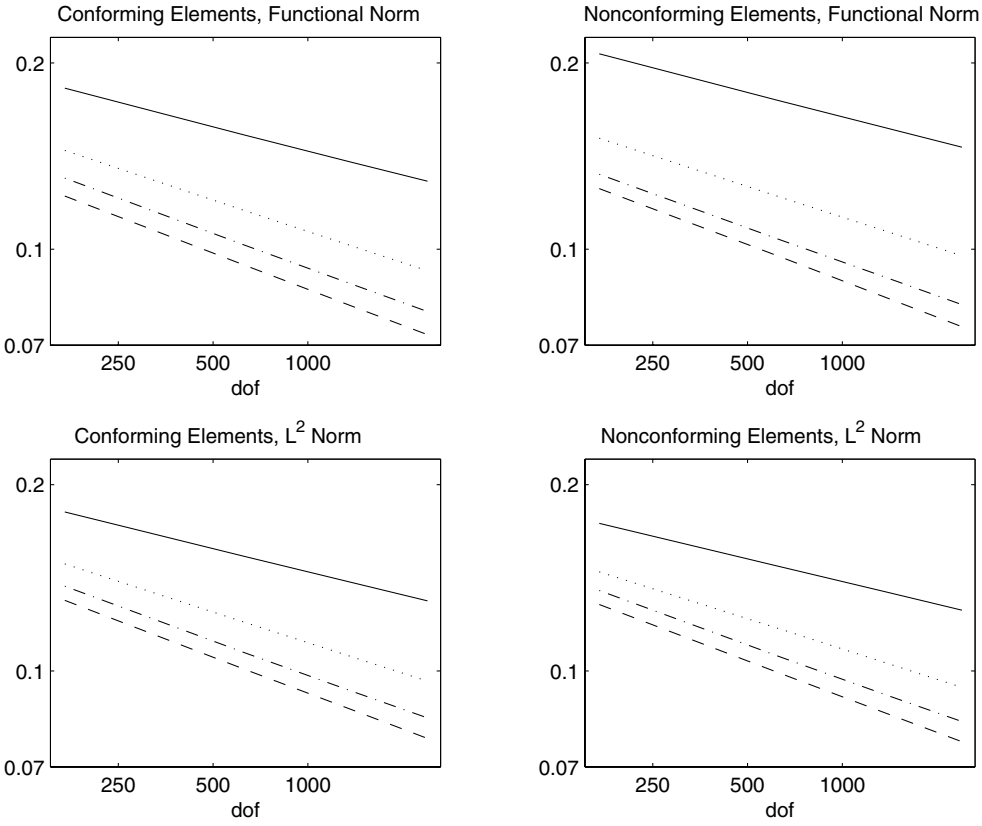

FIG. 2. *Error reductions for (D)LSFEMs of various orders, measured per degree of freedom. For degree k increasing from 1 to 4 (solid, dotted, dash-dotted, dashed), the convergence rate (slope) improves slightly, and higher-order methods exhibit smaller error constants per degree of freedom.*

TABLE 3
Convergence rates for $\theta = \frac{\pi}{8}$ using quadrilaterals and various weights ω.

	$\omega = \frac{1}{h^2}$		$\omega = \frac{1}{h}$		$\omega = 1$		$\omega = h$	
k	L^2	$\mathcal{G}^h$	L^2	$\mathcal{G}^h$	L^2	$\mathcal{G}^h$	L^2	$\mathcal{G}^h$
1	.25	.28	.24	.26	.25	.47	.24	.57
2	.32	.25	.32	.33	.33	.45	.32	.59
3	.36	.37	.36	.34	.38	.44	.37	.52
4	.38	.39	.37	.37	.40	.46	.40	.52

and unacceptable oscillations for many FEMs. However, it was shown in [5, 17] that these negative effects are small in the LS formulation. Overshoots for these solutions are displayed in Figure 5. Even though the LSFEM solutions are not strictly monotone and overshoots and undershoots exist, they are contained in a small region near the discontinuity and do not increase in intensity with increasing polynomial order. Nonconforming elements produced nearly identical (less smooth) oscillation and overshoot profiles; see Figure 5. In Figures 4–5 the number of degrees of freedom for the conforming and nonconforming approximations are within 1% of each other.

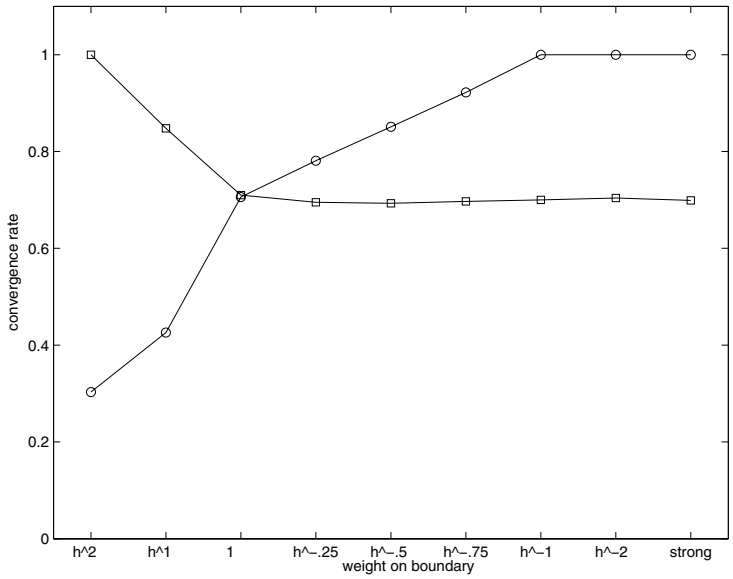

FIG. 3. *Convergence order for the L^2 (squares) and $\mathcal{G}$ (circles) norms as a function of boundary functional weight. For weights stronger than* 1*, the functional error does not converge well. For weights weaker than* 1*, the L^2 error does not converge well. Only for a weight equal to* 1 *are the convergence rates in balance. This agrees with our theoretical results in sections* 2 *and* 3.

TABLE 4
Convergence rates for various θ using quadrilaterals.

		Conforming (4.1)		Nonconforming (4.2) $\omega = \frac{1}{h}$	
θ	k	L^2 norm	$\mathcal{G}$ norm	L^2 norm	$\mathcal{G}^h$ norm
$\frac{\pi}{20}$	1	.25	.25	.25	.23
	2	.33	.33	.33	.32
	3	.35	.35	.35	.35
	4	.36	.35	.37	.35
$\frac{\pi}{12}$	1	.25	.26	.25	.25
	2	.32	.33	.32	.32
	3	.36	.36	.35	.36
	4	.39	.37	.39	.37
$\frac{\pi}{8}$	1	.25	.26	.24	.26
	2	.33	.33	.32	.33
	3	.36	.37	.36	.35
	4	.38	.38	.39	.38
$\frac{\pi}{6}$	1	.25	.26	.24	.26
	2	.33	.34	.32	.33
	3	.37	.37	.37	.37
	4	.39	.39	.39	.39
$\frac{\pi}{4}$	1	.24	.26	.23	.26
	2	.32	.34	.32	.32
	3	.36	.38	.34	.36
	4	.38	.40	.36	.40

TABLE 5
Convergence rates (α) for varying θ using nonconforming linear ($k = 1$) elements on triangles.

θ	0	0.01	0.02	0.03	0.04	0.1
L^2 norm	0.500	0.497	0.485	0.466	0.441	0.304
$\mathcal{G}^h$ norm	0.500	0.498	0.492	0.481	0.468	0.389

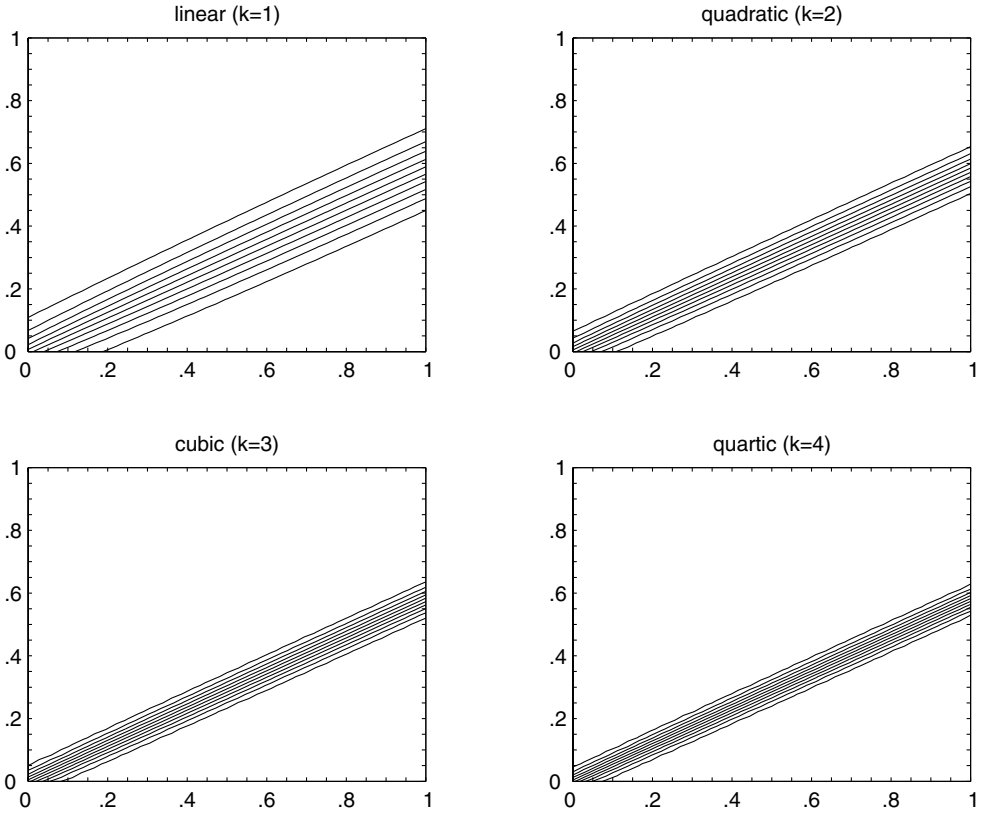

FIG. 4. *Contour plots for various conforming elements. For varying order k, 24, 12, 8, and 6 elements are used in each coordinate direction, respectively.*

7. Multigrid. In this section we address the issues involved in solving the large linear systems arising from the finite element discretizations given in sections 4 and 5. Although the minimization problem is not H^1 equivalent, a property found in many elliptic PDEs that is advantageous for multigrid methods, we will focus on iterative solvers in a multilevel framework employing the techniques of the Ruge–Stüben algebraic multigrid method (AMG) [27]. As we will see, AMG does not fully achieve optimal convergence factors independent of h.

First, consider the problem in the context of the limit case of an anisotropic diffusion operator. More specifically, consider the PDE

$$\begin{aligned} \mathcal{L}_A p &= \tilde{f} \quad \text{on } \Omega, \\ p &= 0 \quad \text{in } \Gamma_I, \\ \mathbf{n} \cdot \nabla p &= 0 \quad \text{in } \Gamma \setminus \Gamma_I, \end{aligned} \tag{7.1}$$

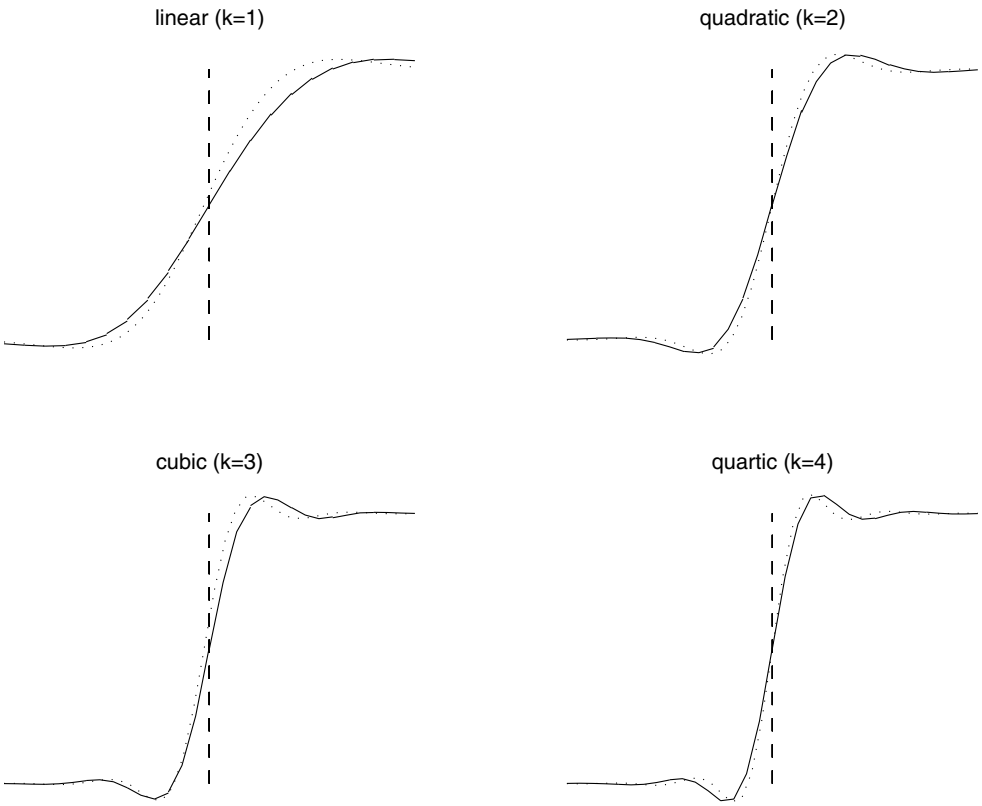

FIG. 5. *Solution profiles for various conforming elements at slice $x = 0.5$. Dotted line: conforming elements. Solid line: nonconforming elements. Dashed line: location of exact discontinuity.*

where

$$\mathcal{L}_A p := \nabla \cdot (A \nabla p), \tag{7.2}$$

and $\tilde{f} \in L^2(\Omega)$.

If $A = I$, we can write $I = \mathbf{b}\mathbf{b}^T + \mathbf{d}\mathbf{d}^T$, where $\mathbf{b} \cdot \mathbf{b} = 1$, $\mathbf{d} \cdot \mathbf{d} = 1$, and $\mathbf{b} \cdot \mathbf{d} = 0$. When $A = \mathbf{b}\mathbf{b}^T + \varepsilon \mathbf{d}\mathbf{d}^T$ for $0 < \varepsilon < 1$, the operator $\mathcal{L}_A$ is an anisotropic diffusion operator because of the strong connection in a particular direction: $\mathbf{b}$. Efficient multigrid algorithms have been developed for this class of PDEs, and we would expect to be able to apply similar algorithms to (1.1)–(1.2).

The Galerkin weak form of (7.1) would be as follows (using standard Sobolev space notation): Find $p \in H^1(\Omega)$ with $p = 0$ on Γ_I such that

$$\langle \mathbf{b} \cdot \nabla p,\, \mathbf{b} \cdot \nabla q \rangle_{0,\Omega} + \varepsilon \langle \mathbf{d} \cdot \nabla p,\, \mathbf{d} \cdot \nabla q \rangle_{0,\Omega} = \langle \tilde{f}, q \rangle_{0,\Omega} \tag{7.3}$$

for every $q \in H^1(\Omega)$ with $q = 0$ on Γ_I. The left-hand side is similar to the left-hand side of the weak form of our LS formulation, which can be written as

$$\langle \mathbf{b} \cdot \nabla p, \mathbf{b} \cdot \nabla q \rangle_{0,\Omega} + \langle p,\, q \rangle_B = \langle f, \mathbf{b} \cdot \nabla q \rangle_{0,\Omega} + \langle g,\, q \rangle_B. \tag{7.4}$$

We test the convergence of AMG using the example flow described in section 6 with $\theta = \frac{\pi}{6}$. Bilinear elements are used for ease of implementation and interpretation

TABLE 6

AMG convergence factors, ρ, for various cycles. w*: weak boundary conditions,* s*: strong boundary conditions.*

$N \times N$	V(1,1)$_{\mathrm{w}}$	V(1,1)$_{\mathrm{s}}$	V(2,2)$_{\mathrm{w}}$	V(2,2)$_{\mathrm{s}}$	W(1,1)$_{\mathrm{w}}$	W(1,1)$_{\mathrm{s}}$
16×16	0.510	0.430	0.420	0.290	0.300	0.250
32×32	0.610	0.500	0.540	0.430	0.350	0.280
64×64	0.700	0.590	0.640	0.530	0.380	0.300
128×128	0.770	0.670	0.730	0.630	0.460	0.300
256×256	0.840	0.740	0.820	0.710	0.530	0.310
512×512	0.910	0.840	0.890	0.810	0.590	0.360

of the work involved. All AMG calculations are done using John Ruge's FOSPACK (first order systems least-squares finite element software package) [26]. The relaxation strategy used in the cycles presented in Table 6 is pointwise Gauss–Seidel. On the downsweep of a cycle, first fine-grid points and then coarse grid points are relaxed, while on the upsweep of a cycle, coarse grid points are relaxed before fine-grid points using a reverse ordering.

The first four convergence columns of Table 6 show the increase in convergence factors in V(1,1) and V(2,2) cycles as the mesh is refined. These values are the factors by which the error is reduced on the finest level by performing one cycle and are the geometric average of convergence factors from one cycle to the next up until the relative residual has reached a prescribed tolerance. We would like these factors to be small—i.e., large reduction in error from one cycle to the next—and we would like these factors to remain constant as the grid is refined. An interesting phenomenon is revealed in the last column of Table 6, which shows W(1,1)-cycle convergence factors with strong treatment of the boundary conditions. The functional for this method is given by the functional in (3.17) without the boundary term $\|u - g\|_B$. This is not the functional we ultimately intend to use, but its resulting linear system exposes some perhaps beneficial aspects of the solver. Notice that the factors are *small*. This is a relative and vague rating, but in the multigrid community grid-independent factors less than 0.5 are generally considered a success. One significant shortcoming of using the implementation of the strong boundary conditions is that $\mathcal{G}(p^h; 0, 0)$ fails to decrease (recall $\mathcal{G}$ is a sharp error estimator). The significance of looking at this case becomes clear when comparing these values to column 6 of Table 6. When we keep the boundary term in the functional (i.e., weak boundary condition), the convergence factors fail to remain constant for the W(1,1)-cycle.

As a measure of grid complexity, we compute the work per cycle in terms of fine-grid relaxation sweeps (or work units) and find growth with n. We compute the complexity by summing the number of nonzero matrix entries on each level multiplied by the number of relaxation sweeps performed on that level, divided by the number of nonzero entries in the fine-grid matrix. This complexity is close to a measure of the work units per cycle.

In Table 7 we report the approximate number of work units per cycle for the tests reported in Table 6. In Table 8 we report the number of work units required to reduce the error by a factor of 10. This "work units per digit" is a measure of the total relative complexity of the algorithm and is computed as

$$W_d = -\frac{W_c}{\log \rho}, \tag{7.5}$$

TABLE 7
Work units per cycle: W_c. w*: weak boundary conditions,* s*: strong boundary conditions.*

$N \times N$	$V(1,1)_w$	$V(1,1)_s$	$V(2,2)_w$	$V(2,2)_s$	$W(1,1)_w$	$W(1,1)_s$
16×16	3.774	3.824	7.549	7.552	5.431	5.553
32×32	4.088	4.098	8.175	8.192	6.890	6.940
64×64	4.241	4.253	8.482	8.504	7.792	7.890
128×128	4.333	4.329	8.667	8.658	8.597	8.418
256×256	4.373	4.373	8.747	8.747	9.101	9.081
512×512	4.396	4.394	8.791	8.787	9.537	9.471

TABLE 8
Work units per digit of accuracy: W_d. w*: weak boundary conditions,* s*: strong boundary conditions.*

$N \times N$	$V(1,1)_w$	$V(1,1)_s$	$V(2,2)_w$	$V(2,2)_s$	$W(1,1)_w$	$W(1,1)_s$
16×16	12.907	10.434	20.036	14.047	10.387	9.223
32×32	19.041	13.613	30.549	22.349	15.112	12.553
64×64	27.379	18.560	43.763	30.844	18.543	15.090
128×128	38.176	24.892	63.410	43.149	25.492	16.099
256×256	57.758	33.443	101.489	58.804	33.007	17.854
512×512	107.321	58.025	173.710	96.020	41.620	21.346

where W_c is the work units per cycle discussed above and ρ is the convergence factor presented in Table 6. Notice in Table 8 that the number of work units per digit for the V(1,1)-cycle appears to be growing slowly with the dimension of the linear system, which increases by a factor of 4 with each row of the table. Likewise, the work units per digit for the W(1,1)-cycle with weak boundary conditions appears to be growing but more slowly, while the work units per digit for the W(1,1)-cycle with strong boundary conditions appears not to grow substantially. Strong boundary conditions do not reduce the growth in complexity with grid size much for V-cycles, while for W(1,1)-cycles the complexity growth is significantly reduced. This shows that W cycles are necessary (more work needs to be done on coarse grids).

8. Conclusion. In this paper we have studied the LSFEM for scalar linear hyperbolic PDEs. We have identified the space of admissible boundary data and have established a Poincaré inequality for the graph norm. We have presented a well-posed formulation of the problem based on the minimization of a LS functional. Finite element solutions were obtained by minimizing the LS functional over a finite dimensional subspace and also by minimizing a similar functional, incorporating a jump term, over a discontinuous nonconforming finite dimensional space. It was also determined that a weight of $\omega \geq c\frac{1}{h}$ was required in order for the uniform Poincaré inequality to hold, where c is a grid-independent constant. Hence, a weight of $\omega = \frac{1}{h}$ was used in the majority of the computational comparisons.

We found, numerically, several advantages in using higher-order elements for discontinuous flow calculation. As the polynomial degree of the finite elements was increased, an increase in convergence rates in the L^2 and the functional norm was observed. The convergence rates were fairly independent of the orientation of the flow, with the exception of *very* small angles, where the convergence rates approached the upper bound of $\frac{1}{2}$—the predicted and confirmed rate for grid-aligned flow. These results were similar for conforming and nonconforming elements and the L^2 and the

functional norms produced nearly identical results. The LS approximations exhibited substantial smearing but only limited oscillation near discontinuities in the solution. Increasing the polynomial degree of the approximation, while keeping the number of degrees of freedom fixed, reduced the smearing with minimal increase in oscillations. There was no apparent advantage of using discontinuous elements over continuous elements for our LS approach. This finding is consistent with the numerical results for DLSFEMs reported in [17]. However, it has to be noted that for other FEMs and in the context of parallelization and locally p-adaptive methods, discontinuous variants may have very important advantages over continuous variants. A good example is the DG method, which has many advantageous properties as described in section 1.

A standard AMG solver based on the Ruge–Stüben algorithm was applied to the linear systems with good results. Nearly grid-independent convergence factors were observed when W-cycles were used and when the boundary conditions were imposed strongly. The relative complexity of this algorithm was nearly independent of the dimension of the linear system. Using the more appropriate formulation involving weak boundary conditions yielded relative complexity that grew slowly with the size of the problem.

While strong enforcement of the boundary conditions will not lead to the solution we seek, the results presented in section 7 suggest a near-optimal numerical scheme for the solution of the linear system that results from using the weak boundary condition. If the value of the approximation at the boundary is known, then the solution of the interior unknowns can be achieved efficiently by solving a system that essentially invokes strong boundary conditions. Thus, a numerical scheme could alternatively solve for the boundary values and then the interior values. We will investigate this approach in a future work.

REFERENCES

[1] R. Abgrall, *Toward the ultimate conservative scheme: following the quest*, J. Comput. Phys., 167 (2001), pp. 277–315.

[2] T. J. Barth, *Numerical methods for gasdynamic systems on unstructured meshes*, in An Introduction to Recent Developments in Theory and Numerics for Conservation Laws (Freiburg/Littenweiler, 1997), Lect. Notes Comput. Sci. Eng. 5, Springer-Verlag, Berlin, 1999, pp. 195–285.

[3] T. J. Barth and H. Deconinck, eds., *High-order methods for computational physics*, Lect. Notes Comput. Sci. Eng. 9, Springer-Verlag, Berlin, 1999.

[4] M. Berndt, T. A. Manteuffel, and S. F. McCormick, *Local error estimates and adaptive refinement for first-order system least squares (FOSLS)*, Electron. Trans. Numer. Anal., 6 (1997), pp. 35–43.

[5] P. B. Bochev and J. Choi, *A comparative study of least-squares, SUPG and Galerkin methods for convection problems*, Int. J. Comput. Fluid Dyn., 15 (2001), pp. 127–146.

[6] P. B. Bochev and J. Choi, *Improved least-squares error estimates for scalar hyperbolic problems*, Comput. Methods Appl. Math., 1 (2001), pp. 115–124.

[7] P. B. Bochev and M. D. Gunzburger, *Finite element methods of least-squares type*, SIAM Rev., 40 (1998), pp. 789–837.

[8] D. Braess, *Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics*, 2nd ed., Cambridge University Press, Cambridge, UK, 2001.

[9] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Texts Appl. Math. 15, Springer-Verlag, New York, 1994.

[10] W. L. Briggs, V. E. Henson, and S. F. McCormick, *A Multigrid Tutorial*, 2nd ed., SIAM, Philadelphia, 2000.

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

[12] Z. Cai, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for second-order partial differential equations: Part* II, SIAM J. Numer. Anal., 34 (1997), pp. 425–454.
[13] G. F. Carey and B. N. Jiang, *Least-squares finite elements for first-order hyperbolic systems*, Internat. J. Numer. Methods Engrg., 26 (1988), pp. 81–93.
[14] B. Cockburn, *Discontinuous Galerkin methods for convection-dominated problems*, in High-Order Methods for Computational Physics, Lect. Notes Comput. Sci. Eng. 9, Springer-Verlag, Berlin, 1999, pp. 69–224.
[15] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Computational Differential Equations*, Cambridge University Press, Cambridge, UK, 1996.
[16] V. Henson and U. M. Yang, *BoomerAMG: A parallel algebraic multigrid solver and preconditioner*, Appl. Numer. Math., 41 (2002), pp. 155–177.
[17] P. Houston, M. Jensen, and E. Süli, *hp-discontinuous Galerkin finite element methods with least-squares stabilization*, J. Sci. Comput., 17 (2002), pp. 3–25.
[18] P. Houston, J. A. Mackenzie, E. Süli, and G. Warnecke, *A posteriori error analysis for numerical approximations of Friedrichs systems*, Numer. Math., 82 (1999), pp. 433–470.
[19] B. N. Jiang, *The Least-Squares Finite Element Method, Theory and Applications in Computational Fluid Dynamics and Electromagnetics*, Scientific Computation, Springer-Verlag, Berlin, 1998.
[20] C. Johnson, U. Nävert, and J. Pitkäranta, *Finite element methods for linear hyperbolic problems*, Comput. Methods Appl. Mech. Engrg., 45 (1984), pp. 285–312.
[21] C. Johnson and J. Pitkäranta, *An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation*, Math. Comp., 46 (1986), pp. 1–26.
[22] R. J. LeVeque, *Numerical Methods for Conservation Laws*, 2nd ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1992.
[23] E. E. Lewis and J. W. F. Miller, *Computational Methods of Neutron Transport*, American Nuclear Society, La Grange Park, IL, 1993.
[24] T. A. Manteuffel, K. J. Ressel, and G. Starke, *A boundary functional for the least-squares finite-element solution of neutron transport problems*, SIAM J. Numer. Anal., 37 (2000), pp. 556–586.
[25] S. F. McCormick, *Multilevel Adaptive Methods for Partial Differential Equations*, Frontiers Appl. Math. 6, SIAM, Philadelphia, 1989.
[26] J. Ruge, *FOSPACK: A First-Order Systems Least-Squares (FOSLS) Code*, manuscript.
[27] J. Ruge and K. Stüben, *Efficient solution of finite difference and finite element equations*, in Multigrid Methods for Integral and Differential Equations (Bristol, 1983), Oxford University Press, New York, 1985, pp. 169–212.
[28] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.
[29] U. Trottenberg, C. W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, San Diego, 2001.
[30] I. Yavneh, C. H. Venner, and A. Brandt, *Fast multigrid solution of the advection problem with closed characteristics*, SIAM J. Sci. Comput., 19 (1998), pp. 111–125.