SIAM J. NUMER. ANAL.
Vol. 32, No. 2, pp. XX–XX, April 1995
© 1995 Society for Industrial and Applied Mathematics
???

A CHARACTERISTICS-MIXED FINITE ELEMENT METHOD FOR ADVECTION DOMINATED TRANSPORT PROBLEMS*

TODD ARBOGAST† AND MARY F. WHEELER†

Abstract. We define a new finite element method, called the characteristics-mixed method, for approximating the solution to an advection dominated transport problem. The method is based on a space-time variational form of the advection-diffusion equation. Our test functions are piecewise constant in space, and in time they approximately follow the characteristics of the advective (i.e., hyperbolic) part of the equation. Thus the scheme uses a characteristic approximation to handle advection in time. This is combined with a low order mixed finite element spatial approximation of the equation. Boundary conditions are incorporated in a natural and mass conservative fashion. The scheme is completely locally conservative; in fact, on the discrete level, fluid is transported along the approximate characteristics. A post-processing step is included in the scheme in which the approximation to the scalar unknown is improved by utilizing the approximate vector flux. This has the effect of improving the rate of convergence of the method. We show that it is optimally convergent to order one in time and at least suboptimally convergent to order 3/2 in space.

Key words. Advection-diffusion equation, characteristics-mixed method, mixed finite element method, characteristics

AMS(MOS) subject classifications. 65M12, 65M25, 65M60, 65B99, 76R05, 76R50, 76S05

1. Introduction.

We consider the model advection-diffusion equation

$$(\phi c)_t + \nabla \cdot (cu - D\nabla c) = f_1 - f_2 c \tag{1.1}$$

for the unknown function $c(x,t)$ in a spatial domain $\Omega \subset \mathbb{R}^d$, $d = 1$, 2, or 3, over a time interval $J = (0,T]$, where subscript t denotes partial differentiation with respect to time. This equation governs such phenomena as the flow of heat within a moving fluid, the transport of dissolved nutrients or contaminants within the groundwater, and the transport of a surfactant or tracer within an incompressible oil in a petroleum reservoir. In the latter two cases, c is the concentration of the miscibly dissolved substance, $\phi(x,t)$ is the porosity of the medium, $u(x,t)$ is the Darcy velocity of the fluid mixture, $D(x,t)$ is the diffusion/dispersion tensor, $f_1(x,t)$ represents the injection wells, and $f_2(x,t) \geq 0$ represents the production wells.

Because of molecular diffusion, D is uniformly positive definite. Although this implies that the equation is uniformly parabolic, in many applications the Peclet number is quite high. Thus advection dominates diffusion, and the equation is nearly hyperbolic in nature. The concentration often develops sharp fronts that are nearly shocks.

It is well-known that strictly parabolic discretization schemes applied to the problem do not work well when it is advection dominated. It is especially difficult to approximate well the sharp fronts and to conserve the material or mass in the system.

Effective discretization schemes recognize to some extent the hyperbolic nature of the equation. Many such schemes have been developed, such as the explicit method

* This work was supported in part by the Department of Energy, the National Science Foundation, and the State of Texas Governor's Energy Office.

† Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77251–1892.

of characteristics, upstream-weighted finite difference schemes [25], interior penalty Galerkin methods [14, 12], higher-order Godunov schemes [9, 2], the streamline diffusion method [19], the modified method of characteristics-Galerkin finite element procedure (MMOC-Galerkin) [16, 18, 13], and the Eulerian-Lagrangian localized adjoint method (ELLAM) [8]. Each method has its advantages and disadvantages. Explicit characteristic and Godunov schemes require that a CFL time step constraint be imposed. Upstream weighting tends to introduce into the solution an excessive amount of numerical diffusion near the sharp fronts. Compared to upstream weighting, the streamline diffusion method reduces the amount of numerical diffusion. It adds a user defined amount biased in the direction of the streamline. The interior penalty Galerkin method is subject to overshoot and undershoot, and although no CFL constraint need be imposed, relatively small time steps must be used in practice. In ELLAM, it can be difficult to evaluate the resulting integrals.

We concentrate on MMOC-Galerkin. It is an implicit scheme, so reasonably large time steps may be used, and it does not numerically diffuse the fronts to a particularly excessive degree. Unfortunately, it has certain inherent difficulties, especially with regard to local mass balance. Since it uses a Galerkin spatial discretization, local constants are not in the space of test functions. As a consequence, there is no discrete, element-by-element mass balance (mass is conserved only globally over all of Ω). It is also difficult to compute the integral of the trace-back concentration, since both the approximate concentration and the test function necessarily vary in space.

In this paper, we propose a new scheme that (theoretically, at least) conserves mass locally. It is similar to MMOC-Galerkin in that we approximate the hyperbolic part of the equation along the characteristics. We use, however, a mixed finite element spatial discretization of the equations. Piecewise constants are then in the set of test functions, so mass is conserved element-by-element. We call our procedure the characteristics-mixed method. It can be viewed as a procedure of ELLAM type.

The origin of our scheme can be seen by considering the advection-diffusion equation in a space-time framework. Choose some domain $R \subset \Omega$ (later considered to be a finite element) and two times $0 \le t_1 < t_2 \le T$. The hyperbolic part of the equation (1.1), $\phi c_t + u \cdot \nabla c$, defines the characteristics $\check{x}(x,t) \in \mathbb{R}^d$ along the interstitial velocity $v = u/\phi$ by

$$\check{x}_t = v(\check{x}, t), \quad t_1 \le t \le t_2, \tag{1.2a}$$

$$\check{x}(x, t_2) = x. \tag{1.2b}$$

Now let $\check{R}(t)$ denote the trace-back of R to time t (see Fig. 1),

$$\check{R}(t) = \{x \in \Omega : x = \check{x}(y,t) \text{ for some } y \in R\},$$

and let $\mathcal{R}$ denote the space-time region that follows the characteristics,

$$\mathcal{R} = \{(x,t) \in \Omega \times J : t_1 \le t \le t_2 \text{ and } x \in \check{R}(t)\}.$$

Also define $\mathcal{B} = \{(x,t) \in \partial\mathcal{R} : x \in \partial\Omega\}$.

Multiply (1.1) by a smooth test function $\psi(x,t)$ and integrate over $\mathcal{R}$. With $\tau(x,t) = \big(u(x,t), \phi(x,t)\big)$ denoting the characteristic direction, the hyperbolic terms

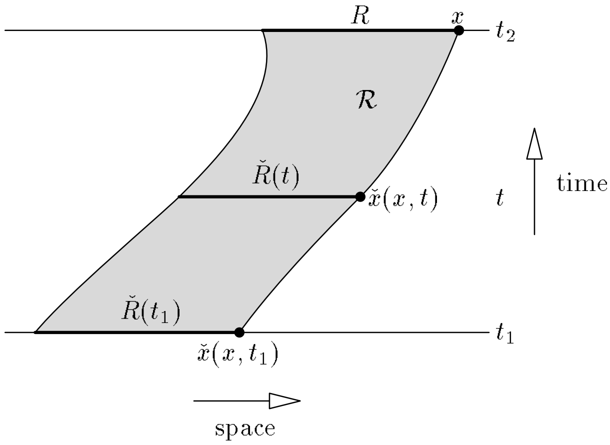

Fig. 1. The space-time region $\mathcal{R}$.

integrate by parts as follows:

$$
\begin{aligned}
&\int_{\mathcal{R}} [\phi c_t + u\cdot\nabla c]\psi\, d\mathcal{R} = \int_{\mathcal{R}} \left(\nabla, \frac{\partial}{\partial t}\right) c\cdot\tau\,\psi\, d\mathcal{R}\\
&\quad = \int_{\partial\mathcal{R}} c\,\tau\cdot\nu_{\mathcal{R}}\,\psi\, d(\partial\mathcal{R}) - \int_{\mathcal{R}} c\left(\nabla,\frac{\partial}{\partial t}\right)\cdot\tau\,\psi\, d\mathcal{R} - \int_{\mathcal{R}} c\,\tau\cdot\left(\nabla,\frac{\partial}{\partial t}\right)\psi\, d\mathcal{R}\\
&\quad = \int_{R}\phi(x,t_2)\,c(x,t_2)\,\psi(x,t_2)\,dR - \int_{\check{R}(t_1)}\phi(x,t_1)\,c(x,t_1)\,\psi(x,t_1)\,d\check{R}(t_1)\\
&\qquad + \int_{\mathcal{B}} c\,u\cdot\nu_{\Omega}\,\psi\, d\mathcal{B} - \int_{\mathcal{R}} c\,(\phi_t + \nabla\cdot u)\psi\, d\mathcal{R} - \int_{\mathcal{R}} c\,(\phi\psi_t + u\cdot\nabla\psi)\, d\mathcal{R},
\end{aligned}
\tag{1.3}
$$

where $\nu_S(x,t)$ is the unit normal to a set S (we used that $\tau\cdot\nu_{\mathcal{R}} = 0$ on the space-time "sides" $\partial\mathcal{R}\cap(t_1,t_2)\setminus\mathcal{B}$). Therefore we have that

$$
\begin{aligned}
&\int_{R}\phi(x,t_2)\,c(x,t_2)\,\psi(x,t_2)\,dR - \int_{\check{R}(t_1)}\phi(x,t_1)\,c(x,t_1)\,\psi(x,t_1)\,d\check{R}(t_1)\\
&\qquad - \int_{\mathcal{R}} c\,(\phi\psi_t + u\cdot\nabla\psi)\, d\mathcal{R} - \int_{\mathcal{R}} \nabla\cdot(D\nabla c)\,\psi\, d\mathcal{R}\\
&\quad = \int_{\mathcal{R}} (f_1 - f_2 c)\,\psi\, d\mathcal{R} - \int_{\mathcal{B}} c\,u\cdot\nu_{\Omega}\,\psi\, d\mathcal{B}.
\end{aligned}
\tag{1.4}
$$

Since the solution is not very smooth in space and time, we concentrate on low order approximations of (1.4). By low order, we mean a mixed method with a piecewise constant approximating space for the unknown scalar function c. Then our test functions ψ are also piecewise constant in space. In time, we let ψ follow the characteristics; that is, for each element R, we have a test function $\psi(x,t)$ that is a constant for $(x,t)\in\mathcal{R}$ and $\psi(x,t) = 0$ elsewhere. The second term on the left-hand side of (1.4) is then easy to compute, since $c(x,t_1)\,\psi(x,t_1)$ is a piecewise polynomial. The third term on the left-hand side vanishes, since ψ trivially satisfies $\phi\psi_t + u\cdot\nabla\psi = 0$. Clearly these are the two critical terms to approximate well.

An outline of the paper follows. In the next section we define an approximation to the characteristics. By considering fluid flow along these *approximate* characteristics, we derive a special, mixed, variational form for our differential problem. This variational form is the basis of our characteristics-mixed method, which we define in Section 3 for the lowest order Raviart-Thomas-Nedelec mixed finite element spaces [22, 21] over fairly general grids.

In Section 4, we present our convergence results. Since the scalar c is approximated by a piecewise constant function, we should expect no better than first order convergence in space and time. However, a post-processing step of our procedure uses the mixed method approximation to the vector $-D\nabla c$ to improve the accuracy of the scalar approximation. As a consequence, we obtain better than first order convergence in space. The error analysis is given in Sections 5-6.

Based on our error analysis, we extract a stability result in Section 7. In Section 8, we remark on the generalization of the characteristics-mixed method to other mixed spaces.

To avoid confusion, the reader should note that in Sections 2 and 3, we define the characteristics-mixed method for a problem with inflow, outflow, and Dirichlet boundary conditions. However, the results of Sections 4-8 assume periodic boundary conditions.

2. Approximate characteristics and the variational problem. We begin this section by defining completely our transport problem. We state it for three boundary conditions, Dirichlet, inflow (Dankwerts or Robin), and outflow (homogeneous Neumann). So let $\partial\Omega = \Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{in}} \cup \Gamma_{\mathrm{out}}$ be decomposed into three disjoint pieces such that $u(x,t)\cdot\nu < 0$ for $x \in \Gamma_{\mathrm{in}}$ and $u(x,t)\cdot\nu \geq 0$ for $x \in \Gamma_{\mathrm{out}}$, where $\nu = \nu_\Omega$. Also let $c_{\mathrm{D}}(x,t)$ denote the Dirichlet concentration and $c_{\mathrm{in}}(x,t)$ denote the inflow concentration. Introducing the diffusive flux z, our transport problem is

$$(\phi c)_t + \nabla\cdot(cu + z) = f_1 - f_2 c \quad \text{in } \Omega \times J, \tag{2.1a}$$

$$z = -D\nabla c \quad \text{in } \Omega \times J, \tag{2.1b}$$

with the boundary conditions

$$c(x,t) = c_{\mathrm{D}}(x,t) \quad \text{on } \Gamma_{\mathrm{D}} \times J, \tag{2.2a}$$

$$z\cdot\nu = \big(c_{\mathrm{in}}(x,t) - c\big)u\cdot\nu \quad \text{on } \Gamma_{\mathrm{in}} \times J, \tag{2.2b}$$

$$z\cdot\nu = 0 \quad \text{on } \Gamma_{\mathrm{out}} \times J, \tag{2.2c}$$

and the initial condition

$$c(x,0) = c^0(x) \quad \text{on } \Omega. \tag{2.3}$$

Assume that our functions are smooth enough for the discussion that follows. Specific assumptions are enumerated in Section 4.

In general, the characteristics can be determined only approximately. For simplicity of discussion, let u be extended as a smooth, bounded vector field outside Ω, and let ϕ be extended smoothly as a uniformly positive function. There are many ways to solve the first order ordinary differential equation (1.2) for the approximate characteristics. We consider here only the Euler method. Alternate schemes can be used, such as the improved Euler or a Runge-Kutta method.

We discretize time by choosing a partition of J, $0 = t^0 < t^1 < \ldots < t^N = T$, and setting $\Delta t^n = t^n - t^{n-1}$. As usual, let $\Delta t = \max_n \Delta t^n$. For a function $\psi(t)$, let $\psi^n = \psi(t^n)$.

The Euler method can be used to solve (1.2) for the approximate characteristics in $M^n \geq 1$ discrete steps as follows: With $t^{n-1,m} = t^{n-1} + m\Delta t^n/M^n$, for $m = M^n, \ldots, 1$, let

$$\check{x}_n^{n-1,m-1}(x) = \check{x}_n^{n-1,m}(x) - v(\check{x}_n^{n-1,m}(x), t^{n-1,m})\Delta t^n/M^n, \tag{2.4a}$$

$$\check{x}_n^{n-1,M^n}(x) = x. \tag{2.4b}$$

Let $\check{x}_n(x,t)$ denote the piecewise linear interpolant in time of the $\check{x}_n^{n-1,m}(x)$. If $\Delta t^n/M^n$ is small enough (depending on the smoothness of v), then the approximate characteristics do not cross each other. We assume this to be the case. We then have a one-to-one mapping $\check{x}_n(\,\cdot\,,t)$ of $\mathbb{R}^d$ into $\mathbb{R}^d$; call its inverse $\hat{x}_n(\,\cdot\,,t)$.

For any time t such that $t^{n-1,m-1} < t \le t^{n-1,m}$, let us define

$$\bar{v}(x,t) = v(\check{x}_n(\hat{x}_n(x,t), t^{n-1,m}), t^{n-1,m}) \quad \text{and} \quad \bar{u} = \bar{v}\phi. \tag{2.5}$$

Our approximate characteristics are defined equivalently with respect to $\bar{v}$:

$$\check{x}_{n,t} = \bar{v}(\check{x}_n, t), \quad t^{n-1} < t < t^n, \tag{2.6a}$$

$$\check{x}_n^n(x) = x. \tag{2.6b}$$

We assume that $\bar{u}\cdot\nu < 0$ on Γ_{in} and $\bar{u}\cdot\nu \ge 0$ on Γ_{out}.

For a function $\psi(x,t)$ and any $t \in J$, if $t^{n-1} < t \le t^n$, let

$$\hat{\psi}(x,t) = \psi(\hat{x}_n(x,t), t^n) \quad \text{and} \quad \check{\psi}(x,t) = \psi(\check{x}_n(x,t), t).$$

Then $\hat{\psi}(x, t^{n-1,+}) = \hat{\psi}^{n-1,+}(x)$ follows the approximate characteristics forward from time t^{n-1} to t^n to become $\psi^n(x)$; this is the type of test function we will use.

The natural, mixed variational problem corresponding to (2.1)–(2.2) is posed in terms of $(c,z) \in L^2(\Omega) \times \mathcal{V}$, where if $H(\Omega;\text{div})$ denotes the set of square integrable vector functions which have a divergence in $L^2(\Omega)$, then $\mathcal{V} = \{V \in H(\Omega;\text{div}) : V\cdot\nu = 0 \text{ on } \Gamma_{\text{out}}\}$. Let us define $\mathcal{W} = \{W \in L^2(\Omega) : W \text{ is piecewise constant}\}$; $\mathcal{W}$ is dense in $L^2(\Omega)$.

We now define our special, mixed, variational form of (2.1)–(2.2). Let $(\,\cdot\,,\,\cdot\,)_S$ denote the $L^2(S)$-inner product (or more generally integration over S), wherein we omit S if $S = \Omega$. Replacing u by $\bar{u} + (u - \bar{u})$ in (2.1a), following the argument given in (1.3)–(1.4) for the velocity $\bar{u}$ and the approximate characteristics $\check{x}_n$, and using (2.2), we obtain for any test function $\psi(x,t) = \hat{W}(x,t)$, $W \in \mathcal{W}$ extended by zero outside Ω, that

$$\begin{aligned}
&(\phi^n c^n, W) - (\phi^{n-1}c^{n-1}, \hat{W}^{n-1,+}) + \int_{t^{n-1}}^{t^n} (\nabla\cdot z, \hat{W})\,dt \\
&\quad = \int_{t^{n-1}}^{t^n} \big[(f_1 - f_2 c, \hat{W}) - (c_{\text{D}}\bar{u}\cdot\nu, \hat{W})_{\Gamma_{\text{D}}} - (c_{\text{in}} u\cdot\nu - z\cdot\nu, \hat{W})_{\Gamma_{\text{in}}}\big]\,dt \\
&\quad\quad + \int_{t^{n-1}}^{t^n} \big\{\big(\nabla\cdot[(\bar{u}-u)c], \hat{W}\big) - \big(c(\bar{u}-u)\cdot\nu, \hat{W}\big)_{\Gamma_{\text{in}}}\big\}\,dt.
\end{aligned} \tag{2.7}$$

This variational equation expresses conservation of mass along the approximate characteristics. Multiply (2.1b) through by $D^{-1}V$, $V \in \mathcal{V}$, and integrate by parts in space to obtain that

$$(D^{-1}z, V) = (c, \nabla\cdot V) - (c, V\cdot\nu)_{\Gamma_{\text{in}}} - (c_{\text{D}}, V\cdot\nu)_{\Gamma_{\text{D}}}. \tag{2.8}$$

It would be difficult to approximate conservatively the inflow boundary conditions in (2.7)–(2.8), since the unknown solution c and z appears in the integrals over Γ_{in}. To rectify this, for $W \in \mathcal{W}$, let $\mathcal{T}(W) = \{R \subset \Omega : W \text{ is constant on } R, R \text{ maximal}\}$, and let $\check{R}_n(t) = \{x \in \Omega : x = \check{x}_n(y,t) \text{ for some } y \in R\} \subset \Omega$. Integrate by parts locally in space on each R the third term on the left-hand side of (2.7) and the first term on the right-hand side of (2.8) to obtain the variational problem

$$(\phi^n c^n, W) - (\phi^{n-1} c^{n-1}, \hat{W}^{n-1,+}) + \int_{t^{n-1}}^{t^n} \sum_{R \in \mathcal{T}(W)} (z \cdot \nu_{\check{R}_n}, \hat{W})_{\partial \check{R}_n \setminus \Gamma_{\rm in}}\, dt$$

$$(2.9a) \qquad = \int_{t^{n-1}}^{t^n} \left[(f_1 - f_2 c, \hat{W}) - (c_{\rm D} \bar{u} \cdot \nu, \hat{W})_{\Gamma_{\rm D}} - (c_{\rm in} u \cdot \nu, \hat{W})_{\Gamma_{\rm in}} \right] dt$$

$$+ \int_{t^{n-1}}^{t^n} \left\{ \left(\nabla \cdot [(\bar{u} - u)c], \hat{W}\right) - \left(c(\bar{u} - u) \cdot \nu, \hat{W}\right)_{\Gamma_{\rm in}} \right\} dt, \quad W \in \mathcal{W},$$

$$(2.9b) \quad (D^{-1} z, V) = \sum_{R \in \mathcal{T}(W)} (c, V \cdot \nu_{\check{R}_n})_{\partial \check{R}_n \setminus \Gamma_{\rm in}} - (c_{\rm D}, V \cdot \nu)_{\Gamma_{\rm D}}, \quad V \in \mathcal{V},\ W \in \mathcal{W}.$$

Although this variational problem uses the "approximate characteristics" defined by $\bar{v}$ (or $\bar{u}$), it is *not* an approximation to the original problem (2.1)–(2.3). Rather, the two are *equivalent* by our derivation using integration by parts in space and time. In fact, any auxiliary function $\bar{u}$ (that does not give rise to crossing characteristics over the time intervals) can be used to produce a variational problem (2.9) equivalent to (2.1)–(2.3). To be useful, however, we need later that $\bar{u} \approx u$ and that the characteristics $\bar{u}$ generates can be computed.

3. The characteristics-mixed method. To discretize space, choose an $h > 0$ and some quasi-uniform grid $\mathcal{T}_h$ over Ω such that the inner and outer diameter of each element is comparable to h. For any element R adjacent to the boundary, we assume that $\partial R \cap \partial \Omega$ is contained in only one of $\Gamma_{\rm D}$, $\Gamma_{\rm in}$, or $\Gamma_{\rm out}$. We also assume that the lowest order Raviart-Thomas-Nedelec mixed finite element spaces [22, 21] $\mathcal{W}_h \times \mathcal{V}_h \subset \mathcal{W} \times \mathcal{V}$ can be defined over our grid (or more generally any similar spaces, e.g., see also [15] for elements with a curved boundary, [24] and [7] for quadrilateral elements, and [10] for prismatic elements). Finally, we need to define a post-processing space $\tilde{\mathcal{W}}_h$ consisting of functions that are discontinuous and piecewise linear over the grid.

We define the characteristics-mixed method from (2.9) as follows. Both $c_h^n \in \mathcal{W}_h$ and $\tilde{c}_h^n \in \tilde{\mathcal{W}}_h$ approximate c^n, and $z_h^n \in \mathcal{V}_h$ approximates z^n. We begin by finding some reasonable approximation $\tilde{c}_h^0 \in \tilde{\mathcal{W}}_h$ to c^0 (such as the $L^2(\Omega)$-projection). For each $n \geq 1$, we define first $(c_h^n, z_h^n) \in \mathcal{W}_h \times \mathcal{V}_h$ such that

$$(\phi^n c_h^n, W) - (\phi^{n-1} \tilde{c}_h^{n-1}, \hat{W}^{n-1,+}) + \sum_{R \in \mathcal{T}_h} (z_h^n \cdot \nu_R, W)_{\partial R \setminus \Gamma_{\rm in}} \Delta t^n$$

$$(3.1a) \qquad = \int_{t^{n-1}}^{t^n} \left[(f_1 - f_2 \hat{c}_h, \hat{W}) - (c_{\rm D} \bar{u} \cdot \nu, \hat{W})_{\Gamma_{\rm D}} - (c_{\rm in} u \cdot \nu, \hat{W})_{\Gamma_{\rm in}} \right] dt, \quad W \in \mathcal{W}_h,$$

$$(3.1b) \qquad ((D^n)^{-1} z_h^n, V) = \sum_{R \in \mathcal{T}_h} (c_h^n, V \cdot \nu_R)_{\partial R \setminus \Gamma_{\rm in}} - (c_{\rm D}^n, V \cdot \nu)_{\Gamma_{\rm D}}, \quad V \in \mathcal{V}_h.$$

We then locally post-process the concentration by finding $\tilde{c}_h^n \in \tilde{\mathcal{W}}_h$ such that on each element $R \in \mathcal{T}_h$,

$$(3.2a)\qquad \big(\phi^n(\tilde{c}_h^n - c_h^n), 1\big)_R = 0,$$

$$(3.2b)\qquad (D^n \nabla \tilde{c}_h^n + z_h^n, \nabla \omega)_R = 0, \quad \omega \in \tilde{\mathcal{W}}_h.$$

If $\Gamma_{\text{in}} = \emptyset$, we can apply the Divergence Theorem to the two terms in (3.1) over $\partial R \setminus \Gamma_{\text{in}}$ to obtain the more usual mixed formulation; (3.1) is written to enable us to handle Γ_{in} in a conservative form. We have also handled the $f_2 c$ term conservatively, and it is computed on an element R as

$$(3.3)\qquad \int_{t^{n-1}}^{t^n} (f_2 \hat{c}_h, \hat{W})_{\check{R}_n(t)}\, dt = (c_h^n W)\big|_R \int_{t^{n-1}}^{t^n} \int_{\check{R}_n(t)} f_2 \, dx\, dt.$$

Mass is conserved locally on each element up to the error in approximating the integrals. In fact, the discrete equations express local conservation of mass in which fluid is transported along the approximate characteristics.

We remark that it is well known that for the mixed method, the error in the approximation of the scalar variable in the $L^2(\Omega)$-norm is only of the first order in h. Our post-processing technique (3.2) is similar to that used by Stenberg for the Stokes problem [23]. It improves the approximation c_h^n so that the error between $\tilde{c}_h^n$ and c^n is of higher order. This post-processing preserves mass on each element, and it is easily computed. It is anti-diffusive, so a slope limiting procedure [20, 11] should be applied to $\tilde{c}_h^n$ to prevent overshoot and undershoot. This has the effect of adding numerical diffusion near sharp fronts, where higher order accuracy cannot be expected.

Issues of implementation will be discussed elsewhere (see also the preliminary report [1]).

LEMMA 1. *If ϕ is uniformly bounded above and below by positive constants, D and D^{-1} are uniformly bounded, symmetric, positive definite tensors, all integrals converge, and the approximate characteristics do not cross, then there exists a unique solution to the characteristics-mixed method.*

Proof. The linear system generated by (3.1)–(3.2) is square, so the existence of a solution is implied by uniqueness. If $\tilde{c}_h^{n-1}$, f_1, c_{D}, and c_{in} are zero, then take $W = c_h^n$, $V = z_h^n$, and $\omega = \tilde{c}_h^n$, to conclude uniqueness, since $f_2 \geq 0$. □

4. Convergence results. We give an analysis of the approximation error in the restricted case that Ω is a rectangular parallelepiped and our problem has periodic boundary conditions. We then have a natural, periodic extension of u and ϕ, and for fixed t, $\check{x}_n(\,\cdot\,, t)$ is known to be a differentiable homeomorphism of Ω to itself, assuming Δt^n is sufficiently small (depending on the smoothness of u and ϕ, see [18] and also Lemma 2 in Section 6). We also assume for convenience that a single Euler step is taken to define the approximate characteristics (i.e., $M^n = 1$ for all n):

$$\check{x}_n(x,t) = x - v(x,t^n)\,(t^n - t) \quad\text{and}\quad \bar{v}(x,t) = v(\hat{x}_n(x,t), t^n) = \hat{v}(x,t).$$

We restate the variational problem (2.9) as

$$(4.1a)\qquad \begin{aligned} &(\phi^n c^n, W) - (\phi^{n-1} c^{n-1}, \hat{W}^{n-1,+}) + \int_{t^{n-1}}^{t^n} (\nabla \cdot z, \hat{W})\, dt \\ &\quad = \int_{t^{n-1}}^{t^n} \big\{ (f_1 - f_2 c, \hat{W}) + \big(\nabla \cdot [(\bar{u} - u)c], \hat{W}\big) \big\}\, dt, \quad W \in \mathcal{W}, \end{aligned}$$

$$(4.1b)\qquad (D^{-1} z, V) = (c, \nabla \cdot V), \quad V \in \mathcal{V},$$

and our characteristics-mixed scheme (3.1) as

$$(\phi^n c_h^n, W) - (\phi^{n-1} \tilde{c}_h^{n-1}, \hat{W}^{n-1,+}) + (\nabla\cdot z_h^n, W)\,\Delta t^n = \int_{t^{n-1}}^{t^n} (f_1 - f_2 \hat{c}_h, \hat{W})\,dt, \quad W \in \mathcal{W}_h, \tag{4.2a}$$

$$\big((D^n)^{-1} z_h^n, V\big) = (c_h^n, \nabla\cdot V), \quad V \in \mathcal{V}_h, \tag{4.2b}$$

together with the post-processing step (3.2), where now $\mathcal{V}_h \subset \mathcal{V} = \{V \in H(\Omega;\mathrm{div}) : V(x)\cdot\nu(x) = V(y)\cdot\nu(x)$whenever $x \in \partial\Omega$ and $y \equiv x$ by periodicity$\}$. Clearly Lemma 1 continues to hold.

We denote by $W^{k,p}(S)$ the standard Sobolev space of k-differentiable functions in $L^p(\Omega)$. Let $\|\cdot\|_{k,S}$ be the norm of $H^k(S) = W^{k,2}(S)$ or $\big(H^k(S)\big)^d$, where we omit S if $S = \Omega$. Let $W^{k,p}(J; W^{j,q}(\Omega))$ denote the usual set of functions with the norm

$$\|\psi\|_{W^{k,p}(J;W^{j,q}(\Omega))} = \left\{ \sum_{i=0}^{k} \int_J \left\| \frac{\partial^i}{\partial t^i}\psi(\,\cdot\,,t)\right\|_{W^{j,q}(\Omega)}^p dt \right\}^{1/p},$$

where if $p=\infty$, the integral is replaced by the essential supremum. We will denote by Q a generic positive constant independent of h, n, the Δt^n, and c.

The standard assumptions are as follows.

(A1) $c, z, \nabla\cdot z \in C^1(J; H^1(\Omega))$.

(A2) $\partial\Omega$ is 2-regular (e.g., $\partial\Omega$ is C^1 or Ω is convex).

(A3) $\phi \in W^{1,\infty}(J; W^{1,\infty}(\Omega))$ and $Q^{-1} \le \phi(x,t)$.

(A4) D is a uniformly bounded, symmetric, positive definite tensor such that $D, D^{-1} \in \big(W^{1,\infty}(J; W^{1,\infty}(\Omega))\big)^{d\times d}$.

(A5) $v, \nabla\cdot v \in W^{1,\infty}(\Omega\times J)$.

(A6) $f_2 \in L^\infty(J; W^{1,\infty}(\Omega))$ and $f_2 \ge 0$.

(A7) $f_1 \in L^1(\Omega\times J)$.

THEOREM 1. *Assume (A1)–(A7). If the initialization error satisfies*

$$\|\tilde{c}_h^0 - c^0\|_0 \le K_0(c^0)\,h^2 \tag{4.3}$$

for some constant K_0 depending on c^0, and if there is some constant $K' > 0$ such that $\Delta t^n \ge K'\,h^{3/2}$, then for h and Δt sufficiently small,

$$\max_n \|\tilde{c}_h^n - c^n\|_0 \le K(c)\,(h^{3/2} + \Delta t), \tag{4.4}$$

$$\max_n \|c_h^n - c^n\|_0 \le K(c)\,(h + \Delta t), \tag{4.5}$$

$$\left\{ \sum_{n=1}^{N} \|z_h^n - z^n\|_0^2\,\Delta t^n \right\}^{1/2} \le K(c)\,(h + \Delta t), \tag{4.6}$$

where K depends on c but not on h or Δt:

$$K(c) = \bar{K}\Bigg\{ \bigg[\int_J \big(\|c\|_1^2 + \|c_\tau\|_0^2 + \|(\nabla\cdot z)_\tau\|_0^2 + \|\nabla\cdot z\|_0^2 + \|z_t\|_1^2 + \|\nabla\cdot z_t\|_1^2 \big)\,dt \bigg]^{1/2} + \max_{t\in J}\big[\,\|z\|_1 + \|\nabla\cdot z\|_1\,\big] + K_0(c^0) \Bigg\}, \tag{4.7}$$

for some constant $\bar{K}$. Moreover, there is some constant $\epsilon' > 0$ such that if $u = 0$ and $\epsilon'\Delta t^n \geq h^2$, then for h and Δt sufficiently small, (4.5)–(4.6) hold and

$$\max_n \|\tilde{c}_h^n - c^n\|_0 \leq K(c)\,(h^2 + \Delta t). \tag{4.8}$$

If there is no advection ($u = 0$), then we have optimal order convergence in space and time; otherwise, we are only able to show a suboptimal result with a loss of 1/2 power on the spatial convergence rate of the post-processed concentration.

In our analysis, we will use the technique of comparing the finite element solution to an elliptic projection $(C_h, Z_h) \in \mathcal{W}_h \times \mathcal{V}_h$ of the true solution (c, z) [26]. Define it by

$$\big(\phi(C_h - c), W\big) + \big(\nabla\cdot(Z_h - z), W\big) = 0, \quad W \in \mathcal{W}_h, \tag{4.9a}$$

$$\big(D^{-1}(Z_h - z), V\big) = (C_h - c, \nabla\cdot V), \quad V \in \mathcal{V}_h. \tag{4.9b}$$

Also define $\tilde{C}_h \in \tilde{\mathcal{W}}_h$ on each element $R \in \mathcal{T}_h$:

$$\big(\phi(\tilde{C}_h - C_h), 1\big)_R = 0, \tag{4.10a}$$

$$(D\nabla\tilde{C}_h + Z_h, \nabla\omega)_R = 0, \quad \omega \in \tilde{\mathcal{W}}_h. \tag{4.10b}$$

THEOREM 2. *Assume (A1)–(A4). Then for each $t \in J$ and for h sufficiently small,*

$$\|C_h - c\|_0 \leq K\|z\|_1\, h, \tag{4.11}$$

$$\|Z_h - z\|_0 \leq K\|z\|_1\, h, \tag{4.12}$$

$$\|\tilde{C}_h - c\|_0 \leq K\{\|z\|_1 + \|\nabla\cdot z\|_1\}\, h^2, \tag{4.13}$$

$$\|(\tilde{C}_h - c)_t\|_0 \leq K\{\|z\|_1 + \|\nabla\cdot z\|_1 + \|z_t\|_1 + \|\nabla\cdot z_t\|_1\} h^2, \tag{4.14}$$

where K is independent of t, c, h, and Δt.

5. Proof of Theorem 2. Results (4.11)–(4.12) are known; moreover, if $\mathcal{P}_{\mathcal{W}_h} : L^2(\Omega) \to \mathcal{W}_h$ denotes the $L^2(\Omega)$-projection operator defined by

$$(\mathcal{P}_{\mathcal{W}_h}\psi - \psi, W) = 0, \quad W \in \mathcal{W}_h,$$

then

$$\|C_h - \mathcal{P}_{\mathcal{W}_h} c\|_0 \leq Q\{\|z\|_1 + \|\nabla\cdot z\|_1\}\, h^2, \tag{5.1}$$

and also

$$\|Z_h - z\|_{-1} \leq Q\{\|z\|_1 + \|\nabla\cdot z\|_1\}\, h^2, \tag{5.2}$$

where $\|\cdot\|_{-1}$ denotes the norm of the dual to $H^1(\Omega)$. These results are derived by making strong use of the Raviart-Thomas Projection operator $\pi_h : \big(H^1(\Omega)\big)^d \to \mathcal{V}_h$, which has the properties that

$$(\nabla\cdot\varphi - \nabla\cdot\pi_h\varphi, W) = 0, \quad W \in \mathcal{W}_h, \tag{5.3}$$

$$\|\varphi - \pi_h\varphi\|_0 \leq Q\|\varphi\|_1\, h. \tag{5.4}$$

Strong use is also made of a duality argument to obtain (5.1)–(5.2). See [15] for a proof in the case of a Dirichlet problem (a straightforward variant of the proof handles our periodic problem–see also the argument given below for obtaining (4.14)).

In general we have only that

$$\|\psi - \mathcal{P}_{\mathcal{W}_h}\psi\|_{L^p(\Omega)} \le Q\|\psi\|_{W^{1,p}(\Omega)}\, h, \quad 1 \le p \le \infty. \tag{5.5}$$

The superconvergence of (5.1) can be exploited to obtain (4.13) as follows.

With

$$\eta = C_h - c \quad \text{and} \quad \tilde\eta = \tilde C_h - c,$$

the post-processing error equation is

$$(\phi\tilde\eta, 1)_R = \big(\phi(C_h - \mathcal{P}_{\mathcal{W}_h}c), 1\big)_R + \big(\phi(\mathcal{P}_{\mathcal{W}_h}c - c), 1\big)_R, \tag{5.6a}$$

$$(D\nabla\tilde\eta, \nabla\omega)_R = (z - Z_h, \nabla\omega)_R, \quad \omega \in \tilde{\mathcal{W}}_h, \tag{5.6b}$$

since $-D\nabla c = z$. For $\tilde\omega \in \tilde{\mathcal{W}}_h$, let $\omega = \tilde\eta + (c - \tilde\omega) \in \tilde{\mathcal{W}}_h$ to obtain that

$$\begin{aligned}(D\nabla\tilde\eta, \nabla\tilde\eta)_R &= \big(z - Z_h, \nabla\tilde\eta + \nabla(c - \tilde\omega)\big)_R - \big(D\nabla\tilde\eta, \nabla(c - \tilde\omega)\big)_R \\ &\le Q\big\{\|z - Z_h\|_{0,R}^2 + \|\nabla(c - \tilde\omega)\|_{0,R}^2\big\} + \tfrac12\|D^{1/2}\nabla\tilde\eta\|_{0,R}^2.\end{aligned}$$

Since $\tilde{\mathcal{W}}_h$ contains the linear functions, the Bramble-Hilbert Lemma [3, 4, 17], a good choice of $\tilde\omega$, and (4.12) imply that

$$\Big\{\sum_{R\in\mathcal{T}_h} \|\nabla\tilde\eta\|_{0,R}^2\Big\}^{1/2} \le Q\|z\|_1\, h. \tag{5.7}$$

Now for any constant W, by (5.6a),

$$(\phi\tilde\eta, \tilde\eta)_R = (\phi\tilde\eta, \tilde\eta - W)_R + \big(\phi(C_h - \mathcal{P}_{\mathcal{W}_h}c), W\big)_R + \big(\phi(\mathcal{P}_{\mathcal{W}_h}c - c), W\big)_R,$$

and this last term is

$$\big(\phi(\mathcal{P}_{\mathcal{W}_h}c - c), W\big)_R = \big((\phi - \mathcal{P}_{\mathcal{W}_h}\phi)(\mathcal{P}_{\mathcal{W}_h}c - c), W\big)_R.$$

We therefore estimate that

$$\begin{aligned}\|\tilde\eta\|_{0,R}^2 \le Q\big\{&\|\tilde\eta - W\|_{0,R}^2 + \|C_h - \mathcal{P}_{\mathcal{W}_h}c\|_{0,R}^2 \\ &+ \|\phi - \mathcal{P}_{\mathcal{W}_h}\phi\|_{L^\infty(R)}^2\|\mathcal{P}_{\mathcal{W}_h}c - c\|_{0,R}^2\big\}.\end{aligned}$$

Since a good choice of W implies that

$$\|\tilde\eta - W\|_{0,R} \le Q\|\nabla\tilde\eta\|_{0,R}\, h,$$

we obtain (4.13) with (5.7), (5.1), and (5.5).

To prove (4.14), differentiate (4.9) in time:

$$(\phi\eta_t, W) + \big(\nabla\cdot(Z_h - z)_t, W\big) = -(\phi_t\eta, W), \quad W \in \mathcal{W}_h, \tag{5.8a}$$

$$\big(D^{-1}(Z_h - z)_t, V\big) = (\eta_t, \nabla\cdot V) - \big((D^{-1})_t(Z_h - z), V\big), \quad V \in \mathcal{V}_h. \tag{5.8b}$$

Using (4.11)–(4.12), an analysis as in [15] leads to

$$\|\eta_t\|_0 \le Q\{\|z\|_1 + \|z_t\|_1\}h, \tag{5.9}$$

$$\|(Z_h - z)_t\|_0 \le Q\{\|z\|_1 + \|z_t\|_1\}h, \tag{5.10}$$

$$\|\nabla\cdot(Z_h - z)_t\|_0 \le Q\{\|z\|_1 + \|z_t\|_1 + \|\nabla\cdot z_t\|_1\}h. \tag{5.11}$$

In fact, superconvergence is obtained for $(C_h - \mathcal{P}_{\mathcal{W}_h} c)_t$. Since this result is critical, we derive it in detail.

Let ψ solve the elliptic problem

$$\begin{aligned} &\phi\psi - \nabla\cdot(D\nabla\psi) = (C_h - \mathcal{P}_{\mathcal{W}_h} c)_t \quad \text{in } \Omega, \\ &\psi = 0 \quad \text{on } \partial\Omega, \end{aligned}$$

so that

$$\|\psi\|_2 \le Q\|(C_h - \mathcal{P}_{\mathcal{W}_h} c)_t\|_0 \tag{5.12}$$

and

$$\|(C_h - \mathcal{P}_{\mathcal{W}_h} c)_t\|_0^2 = \big(\phi(C_h - \mathcal{P}_{\mathcal{W}_h} c)_t, \psi\big) - \big((C_h - \mathcal{P}_{\mathcal{W}_h} c)_t, \nabla\cdot(D\nabla\psi)\big). \tag{5.13}$$

The last term above can be expanded by using (5.8b) with $V = \pi_h(D\nabla\psi)$:

$$\begin{aligned} &-\big((C_h - \mathcal{P}_{\mathcal{W}_h} c)_t, \nabla\cdot(D\nabla\psi)\big) \\ &\quad = -\big(\mathcal{P}_{\mathcal{W}_h}\eta_t, \nabla\cdot(D\nabla\psi)\big) \\ &\quad = -\big(\mathcal{P}_{\mathcal{W}_h}\eta_t, \nabla\cdot\pi_h(D\nabla\psi)\big) \\ &\quad = -\big(\eta_t, \nabla\cdot\pi_h(D\nabla\psi)\big) \\ &\quad = -\big(D^{-1}(Z_h - z)_t, \pi_h(D\nabla\psi)\big) - \big((D^{-1})_t(Z_h - z), \pi_h(D\nabla\psi)\big) \\ &\quad = -\big(D^{-1}(Z_h - z)_t, \pi_h(D\nabla\psi) - D\nabla\psi\big) - \big((Z_h - z)_t, \nabla\psi\big) \\ &\qquad - \big((D^{-1})_t(Z_h - z), \pi_h(D\nabla\psi) - D\nabla\psi\big) - \big((D^{-1})_t(Z_h - z), D\nabla\psi\big). \end{aligned} \tag{5.14}$$

The second term on the far right-hand side above can be integrated by parts and combined with (5.8a) for any $W \in \mathcal{W}_h$ to obtain

$$\begin{aligned} -\big((Z_h - z)_t, \nabla\psi\big) &= \big(\nabla\cdot(Z_h - z)_t, \psi\big) \\ &= \big(\nabla\cdot(Z_h - z)_t, \psi - W\big) - (\phi\eta_t, W) - (\phi_t\eta, W). \end{aligned} \tag{5.15}$$

The next to last term above is

$$\begin{aligned} -(\phi\eta_t, W) &= -\big(\phi(C_h - \mathcal{P}_{\mathcal{W}_h} c)_t, W\big) - \big(\phi(\mathcal{P}_{\mathcal{W}_h} c - c)_t, W\big) \\ &= -\big(\phi(C_h - \mathcal{P}_{\mathcal{W}_h} c)_t, W\big) - \big((\phi - \mathcal{P}_{\mathcal{W}_h}\phi)(\mathcal{P}_{\mathcal{W}_h} c - c)_t, W\big). \end{aligned} \tag{5.16}$$

The last term in (5.15) is similarly expanded. Note that we can combine one term in

(5.16) with (5.13), so (5.13)–(5.16) yields that

$$
\begin{aligned}
&\|(C_h - \mathcal{P}_{\mathcal{W}_h} c)_t\|_0^2 \\
&\quad = \big(\phi(C_h - \mathcal{P}_{\mathcal{W}_h} c)_t, \psi - W\big) - \big(D^{-1}(Z_h - z)_t, \pi_h(D\nabla\psi) - D\nabla\psi\big) \\
&\qquad - \big((D^{-1})_t(Z_h - z), \pi_h(D\nabla\psi) - D\nabla\psi\big) - \big((D^{-1})_t(Z_h - z), D\nabla\psi\big) \\
&\qquad + \big(\nabla\cdot(Z_h - z)_t, \psi - W\big) - \big((\phi - \mathcal{P}_{\mathcal{W}_h}\phi)(\mathcal{P}_{\mathcal{W}_h} c - c)_t, W\big) \\
&\qquad - \big(\phi_t(C_h - \mathcal{P}_{\mathcal{W}_h} c), W\big) - \big((\phi - \mathcal{P}_{\mathcal{W}_h}\phi)_t(\mathcal{P}_{\mathcal{W}_h} c - c), W\big) \\
&\quad \le Q\big\{\|(C_h - \mathcal{P}_{\mathcal{W}_h} c)_t\|_0\|\psi - W\|_0 \\
&\qquad + \big[\|(Z_h - z)_t\|_0 + \|Z_h - z\|_0\big]\|\pi_h(D\nabla\psi) - D\nabla\psi\|_0 \\
&\qquad + \|Z_h - z\|_{-1}\|\psi\|_2 + \|\nabla\cdot(Z_h - z)_t\|_0\|\psi - W\|_0 \\
&\qquad + \big[\|\phi - \mathcal{P}_{\mathcal{W}_h}\phi\|_{L^\infty(\Omega)}\|(\mathcal{P}_{\mathcal{W}_h} c - c)_t\|_0 + \|C_h - \mathcal{P}_{\mathcal{W}_h} c\|_0 \\
&\qquad\quad + \|(\phi - \mathcal{P}_{\mathcal{W}_h}\phi)_t\|_{L^\infty(\Omega)}\|\mathcal{P}_{\mathcal{W}_h} c - c\|_0\big]\|W\|_0\big\} \\
&\quad \le Q\big\{\big[\|(C_h - \mathcal{P}_{\mathcal{W}_h} c)_t\|_0 + \|(Z_h - z)_t\|_0 + \|Z_h - z\|_0\big] h \\
&\qquad + \|Z_h - z\|_{-1} + \|\nabla\cdot(Z_h - z)_t\|_0\, h \\
&\qquad + \|\nabla c_t\|_0\, h^2 + \|C_h - \mathcal{P}_{\mathcal{W}_h} c\|_0 + \|\nabla c\|_0\, h^2\big\}\|\psi\|_2 \\
&\quad \le Q\big\{\|(C_h - \mathcal{P}_{\mathcal{W}_h} c)_t\|_0\, h \\
&\qquad + \big[\|z\|_1 + \|\nabla\cdot z\|_1 + \|z_t\|_1 + \|\nabla\cdot z_t\|_1\big]\, h^2\big\}\|(C_h - \mathcal{P}_{\mathcal{W}_h} c)_t\|_0,
\end{aligned} \tag{5.17}
$$

wherein for the second inequality we use (5.4), (5.5), and the fact that a good choice of W gives us that

$$\|\psi - W\|_0 \le Q\|\psi\|_1\, h \quad \text{and} \quad \|W\|_0 \le Q\{\|\psi\|_0 + \|\psi\|_1\, h\},$$

and also wherein (5.17) for the third inequality we use the results (4.12), (5.1)–(5.2), and (5.10)–(5.12). For h sufficiently small, we obtain our desired superconvergence

$$\|(C_h - \mathcal{P}_{\mathcal{W}_h} c)_t\|_0 \le Q\{\|z\|_1 + \|\nabla\cdot z\|_1 + \|z_t\|_1 + \|\nabla\cdot z_t\|_1\}h^2. \tag{5.18}$$

As in the proof of (4.13), (5.1) and (5.18) can be exploited in an analysis of the derivative of (5.6) to obtain (4.14).

6. Proof of Theorem 1. Let us define

$$\xi^n = c_h^n - C_h^n \in \mathcal{W}_h, \quad \tilde{\xi}^n = \tilde{c}_h^n - \tilde{C}_h^n \in \tilde{\mathcal{W}}_h, \quad \text{and} \quad \zeta^n = z_h^n - Z_h^n \in \mathcal{V}_h.$$

Combining (4.1), (4.2), and (4.9), the error equation can be written as

$$
\begin{aligned}
&(\phi^n(\xi^n+\eta^n),W)-(\phi^{n-1}(\tilde\xi^{n-1}+\tilde\eta^{n-1}),\hat W^{n-1,+})\\
&\qquad+(\nabla\cdot\zeta^n,W)\,\Delta t^n+\int_{t^{n-1}}^{t^n}(f_2\hat\xi,\hat W)\,dt\\
&\quad=\int_{t^{n-1}}^{t^n}(f_2(c-\hat C_h),\hat W)\,dt-\int_{t^{n-1}}^{t^n}\big(\nabla\cdot[(\bar u-u)c],\hat W\big)\,dt \qquad (6.1a)\\
&\qquad+\int_{t^{n-1}}^{t^n}(\nabla\cdot z,\hat W)\,dt-(\nabla\cdot z^n,W)\,\Delta t^n+(\phi^n\eta^n,W)\,\Delta t^n,\\
&\qquad\qquad W\in\mathcal{W}_h,
\end{aligned}
$$

$$
\big((D^n)^{-1}\zeta^n,V\big)=(\xi^n,\nabla\cdot V),\quad V\in\mathcal{V}_h. \qquad (6.1b)
$$

Take $W=\xi^n$ and $V=\zeta^n$ in (6.1), add the two equations, and use (3.2a) and (4.10a) to obtain that

$$
\begin{aligned}
&(\phi^n\tilde\xi^n,\xi^n)-(\phi^{n-1}\tilde\xi^{n-1},\hat\xi^{n-1,+})\\
&\qquad+\big((D^n)^{-1}\zeta^n,\zeta^n\big)\,\Delta t^n+\int_{t^{n-1}}^{t^n}(f_2\hat\xi,\hat\xi)\,dt\\
&\quad=\int_{t^{n-1}}^{t^n}\big(f_2(c-\hat C_h),\hat\xi\big)\,dt-\int_{t^{n-1}}^{t^n}\big(\nabla\cdot[(\bar u-u)c],\hat\xi\big)\,dt \qquad (6.2)\\
&\qquad+\int_{t^{n-1}}^{t^n}\big[(\nabla\cdot z,\hat\xi)-(\nabla\cdot z^n,\xi^n)\big]\,dt\\
&\qquad-(\phi^n\tilde\eta^n,\xi^n)(1-\Delta t^n)+(\phi^{n-1}\tilde\eta^{n-1},\hat\xi^{n-1,+}).
\end{aligned}
$$

The first three terms on the right-hand side above represent primarily time discretization errors. We will estimate them sequentially. It will be helpful to note the following general results.

LEMMA 2. *If $\psi_1\in L^1(J;L^p(\Omega))$ and $\psi_2\in L^{p'}(\Omega)$, $1\le p\le\infty$, $\frac{1}{p}+\frac{1}{p'}=1$, then for almost every fixed t such that $t^{n-1}<t\le t^n$,*

$$
(\psi_1,\hat\psi_2)=(\check\psi_1,\psi_2\mathcal{J}_n)=(\check\psi_1,\psi_2)+(\check\psi_1,\psi_2\mathcal{J}_n') \qquad (6.3)
$$

for some $\mathcal{J}_n(x,t)=1+\mathcal{J}_n'(x,t)$ such that

$$
|\mathcal{J}_n'(x,t)|\le K(t^n-t), \qquad (6.4)
$$

where K depends on $\|v\|_{L^\infty(J;W^{1,\infty}(\Omega))}$. Moreover, for some positive K_1 and K_2 depending on K, and for Δt^n sufficiently small,

$$
\big(1+K_1(\Delta t^n)^{1/p}\big)\|\psi_1\|_{L^p(\Omega)}\le\|\check\psi_1\|_{L^p(\Omega)}\le\big(1+K_2(\Delta t^n)^{1/p}\big)\|\psi_1\|_{L^p(\Omega)}, \qquad (6.5)
$$

$$
\big(1+K_1(\Delta t^n)^{1/p'}\big)\|\psi_2\|_{L^{p'}(\Omega)}\le\|\hat\psi_2\|_{L^{p'}(\Omega)}\le\big(1+K_2(\Delta t^n)^{1/p'}\big)\|\psi_2\|_{L^{p'}(\Omega)}. \qquad (6.6)
$$

proof. For any coordinates i and j,

$$
\frac{\partial\check x_{n,i}(x,t)}{\partial x_j}=\delta_{ij}-\frac{\partial v_i(x,t^n)}{\partial x_j}(t^n-t), \qquad (6.7)
$$

where δ_{ij} is the Kronecker delta. The change of variables $x\mapsto\check x_n(x,t)$ gives (6.3), where $\mathcal{J}_n(x,t)$ is the determinant of the Jacobian $\partial \check{x}_n(x,t)/\partial x$. It has the form

$$\mathcal{J}_n = 1 - \nabla \cdot v^n\,(t^n - t) + \sum_{\ell=2}^{d} \mathcal{R}_n^\ell (t^n - t)^\ell, \tag{6.8}$$

where the $\mathcal{R}_n^\ell$ depend on products of derivatives of v^n; thus, (6.4) holds. Provided that $K\Delta t^n$ is strictly less than one, the final two results follow from (6.3)–(6.4) (unless p or p' is ∞, but these cases are trivial). □

LEMMA 3. *If $\psi \in W^{1,p}(\Omega \times J)$, $1 \le p \le \infty$, $t^{n-1} < t \le t^n$, and ψ_τ is defined by*

$$\begin{aligned}\psi_\tau(\check{x}_n(x,t),t) &= \frac{\partial}{\partial t}\big(\check{\psi}(x,t)\big) = \frac{\partial}{\partial t}\big(\psi(\check{x}_n(x,t),t)\big)\\ &= \psi_t(\check{x}_n(x,t),t) + \bar{v}(\check{x}_n(x,t),t)\cdot\nabla\psi(\check{x}_n(x,t),t)\end{aligned} \tag{6.9}$$

as the derivative along the approximate characteristic (recall (2.6)), then

$$\psi^n(x) - \check{\psi}(x,t) = \int_t^{t^n} \psi_\tau(\check{x}_n(x,s),s)\,ds \tag{6.10}$$

and, for Δt^n sufficiently small,

$$\|\psi^n - \check{\psi}\|_{L^p(\Omega)} \le \big(1 + K(\Delta t^n)^{1/p}\big)\|\psi_\tau\|_{L^p(\Omega\times(t^{n-1},t^n))}\,(\Delta t^n)^{1-1/p}, \tag{6.11}$$

$$\|\hat{\psi} - \psi\|_{L^p(\Omega)} \le \big(1 + K(\Delta t^n)^{1/p}\big)\|\psi_\tau\|_{L^p(\Omega\times(t^{n-1},t^n))}\,(\Delta t^n)^{1-1/p}, \tag{6.12}$$

where K is a constant depending on $\|v\|_{L^\infty(J;W^{1,\infty}(\Omega))}$.

proof. Result (6.10) is immediate, and (6.11) follows with the Cauchy-Schwarz inequality and (6.5). Since $\check{\hat{\psi}} = \psi^n$, (6.12) follows from (6.5) and (6.11). □

LEMMA 4. *If $\psi \in L^2(\Omega)$ and $a \in W^{1,\infty}(\Omega \times J)$, then for any n,*

$$(a^{n-1}\hat{\psi}^{n-1,+}, \hat{\psi}^{n-1,+}) \le (a^n\psi,\psi) + K\|\psi\|_0^2\,\Delta t^n, \tag{6.13}$$

where $K \ge 0$ depends on $\|v\|_{L^\infty(J;W^{1,\infty}(\Omega))}$ and $\|a\|_{W^{1,\infty}(\Omega\times J)}$.

Proof. Use Lemma 2 to make the change of variables

$$\begin{aligned}(a^{n-1}\hat{\psi}^{n-1,+}, \hat{\psi}^{n-1,+}) &= (\check{a}^{n-1,+}\psi, \psi\mathcal{J}_n)\\ &= (a^n\psi,\psi) - \big((a^n - \check{a}^{n-1,+})\psi,\psi\big) + (\check{a}^{n-1,+}\psi, \psi\mathcal{J}_n')\\ &\le (a^n\psi,\psi) + \{\|a^n - \check{a}^{n-1,+}\|_{L^\infty(\Omega)} + Q\|a^{n-1}\|_{L^\infty(\Omega)}\Delta t^n\}\|\psi\|_0^2,\end{aligned}$$

and then apply (6.11). □

Returning to the right-hand side of (6.2), by Lemma 2, the first term is

$$\begin{aligned}\int_{t^{n-1}}^{t^n}\big(f_2(c-\hat{C}_h),\hat{\xi}\big)\,dt &= \int_{t^{n-1}}^{t^n}\big(\check{f}_2\,(\check{c}-C_h^n),\xi^n\mathcal{J}_n\big)\,dt\\ &= \int_{t^{n-1}}^{t^n}\big(\check{f}_2\,(c^n-C_h^n),\xi^n\mathcal{J}_n\big)\,dt + \int_{t^{n-1}}^{t^n}\big(\check{f}_2\,(\check{c}-c^n),\xi^n\mathcal{J}_n\big)\,dt\\ &\le -\int_{t^{n-1}}^{t^n}\big(\check{f}_2\,\eta^n,\xi^n\big)\,dt\\ &\quad + Q\Big\{\|\eta^n\|_0^2(\Delta t^n)^2 + \|\xi^n\|_0^2\Delta t^n + \int_{t^{n-1}}^{t^n}\|\check{c}-c^n\|_0^2\,dt\Big\},\end{aligned} \tag{6.14}$$

where Q is a generic positive constant that depends on $\|v\|_{L^\infty(J;W^{1,\infty}(\Omega))}$ and also on $\|f_2\|_{L^\infty(J;L^\infty(\Omega))}$. (In this section, we will keep track only of new quantities that Q

may depend on and assume that it depends on all previous quantities.) Using (4.10a),

$$\begin{aligned} -(\check{f}_2\,\eta^n,\xi^n) &= \Big(\phi\Big[\mathcal{P}_{\mathcal{W}_h}\Big(\frac{\check{f}_2}{\phi}\Big)-\frac{\check{f}_2}{\phi}\Big]\eta^n,\xi^n\Big)-\Big(\phi\mathcal{P}_{\mathcal{W}_h}\Big(\frac{\check{f}_2}{\phi}\Big)\tilde{\eta}^n,\xi^n\Big)\\ &\le Q\{\|\eta^n\|_0\,h+\|\tilde{\eta}^n\|_0\}\|\xi^n\|_0, \end{aligned} \tag{6.15}$$

where Q depends also on $\|\phi\|_{L^\infty(J;W^{1,\infty}(\Omega))}$, $\|f_2\|_{L^\infty(J;W^{1,\infty}(\Omega))}$, and on the lower positive bound for ϕ. Therefore with Lemma 3,

$$\begin{aligned} &\int_{t^{n-1}}^{t^n}\big(f_2(c-\hat{C}_h),\hat{\xi}\big)\,dt\\ &\quad\le Q\Big\{\|\eta^n\|_0^2\,(h^2+\Delta t^n)+\|\tilde{\eta}^n\|_0^2+\|\xi^n\|_0^2+\int_{t^{n-1}}^{t^n}\|c_\tau\|_0^2\,dt\,\Delta t^n\Big\}\Delta t^n. \end{aligned} \tag{6.16}$$

For the second term on the right-hand side of (6.2),

$$\begin{aligned} &-\int_{t^{n-1}}^{t^n}\big(\nabla\cdot[(\bar{u}-u)c],\hat{\xi}\big)\,dt\\ &\qquad\le Q\Big\{\|\xi^n\|_0^2\Delta t^n+\int_{t^{n-1}}^{t^n}\|\nabla\cdot[(\bar{u}-u)c]\|_0^2\,dt\Big\}. \end{aligned} \tag{6.17}$$

Note that

$$\nabla\cdot[(\bar{u}-u)c]=\nabla\cdot\big((\bar{v}-v)c\phi\big)=\nabla\cdot(\bar{v}-v)\,c\phi+(\bar{v}-v)\cdot\nabla(c\phi). \tag{6.18}$$

The chain rule and (6.7) show that

$$\begin{aligned} \nabla\cdot\bar{v}(x,t)&=\nabla\cdot\big(v(\hat{x}_n(x,t),t^n)\big)\\ &=\nabla\cdot v(\hat{x}_n(x,t),t^n)+\mathcal{O}\big((t^n-t)\big)\\ &=(\widehat{\nabla\cdot v})(x,t)+\mathcal{O}\big((t^n-t)\big), \end{aligned} \tag{6.19}$$

where the last term depends only on $\|v\|_{L^\infty(J;W^{1,\infty}(\Omega))}$, so

$$\begin{aligned} &\int_{t^{n-1}}^{t^n}\|\nabla\cdot(\bar{v}-v)\,c\phi\|_0^2\,dt\\ &\quad\le Q\{\|(\widehat{\nabla\cdot v})-\nabla\cdot v\|_{L^\infty(\Omega\times J)}^2+(\Delta t^n)^2\}\int_{t^{n-1}}^{t^n}\|c\|_0^2\,dt\\ &\quad\le Q\int_{t^{n-1}}^{t^n}\|c\|_0^2\,dt\,(\Delta t^n)^2, \end{aligned} \tag{6.20}$$

where Q depends on $\|(\nabla\cdot v)_\tau\|_{L^\infty(\Omega\times J)}$ by Lemma 3. Also,

$$\bar{v}(x,t)=\hat{v}(x,t)=v(x,t)+\int_t^{t^n}v_t(x,s)\,ds-\int_t^{t^n}\frac{\partial}{\partial s}\hat{v}(x,s)\,ds. \tag{6.21}$$

Therefore,

$$-\int_{t^{n-1}}^{t^n}\big(\nabla\cdot[(\bar{u}-u)c],\hat{\xi}\big)\,dt\le Q\Big\{\|\xi^n\|_0^2\Delta t^n+\int_{t^{n-1}}^{t^n}\|c\|_1^2\,dt\,(\Delta t^n)^2\Big\}, \tag{6.22}$$

where Q depends also on $\|v\|_{W^{1,\infty}(\Omega\times J)}$.

The third term on the right-hand side of (6.2) is estimated with Lemmas 2 and 3:

$$
\begin{aligned}
&\int_{t^{n-1}}^{t^n} \{(\nabla\cdot z,\hat{\xi}) - (\nabla\cdot z^n,\xi^n)\}\,dt \\
&\qquad\le Q\Big\{\|\xi^n\|_0^2\Delta t^n + \int_{t^{n-1}}^{t^n} \|(\nabla\check{\cdot} z)\mathcal{J}_n - \nabla\cdot z^n\|_0^2\,dt\Big\} \\
&\qquad\le Q\Big\{\|\xi^n\|_0^2\Delta t^n + \int_{t^{n-1}}^{t^n} \big[\|(\nabla\cdot z)_\tau\|_0^2 + \|\nabla\cdot z\|_0^2\big]\,dt\,(\Delta t^n)^2\Big\}.
\end{aligned}
\tag{6.23}
$$

Combining (6.2) with (6.16), (6.22), and (6.23) yields the estimate

$$
\begin{aligned}
&(\phi^n\tilde{\xi}^n,\xi^n) - (\phi^{n-1}\tilde{\xi}^{n-1},\hat{\xi}^{n-1,+}) \\
&\qquad + \big((D^n)^{-1}\zeta^n,\zeta^n\big)\,\Delta t^n + \int_{t^{n-1}}^{t^n} (f_2\hat{\xi},\hat{\xi})\,dt \\
&\quad\le Q\Big\{\|\xi^n\|_0^2 + \|\eta^n\|_0^2(h^2+\Delta t^n) + \|\tilde{\eta}^n\|_0^2 + \int_{t^{n-1}}^{t^n}\big[\|c\|_1^2 + \|c_\tau\|_0^2\big]\,dt\,\Delta t^n \\
&\qquad + \int_{t^{n-1}}^{t^n}\big[\|(\nabla\cdot z)_\tau\|_0^2 + \|\nabla\cdot z\|_0^2\big]\,dt\,\Delta t^n\Big\}\Delta t^n \\
&\qquad - (\phi^n\tilde{\eta}^n,\xi^n) + (\phi^{n-1}\tilde{\eta}^{n-1},\hat{\xi}^{n-1,+}).
\end{aligned}
\tag{6.24}
$$

We now analyze the first two terms on the left-hand side of (6.24). They nearly collapse under summation. Use Lemma 4 to estimate that

$$
\begin{aligned}
&(\phi^{n-1}\tilde{\xi}^{n-1},\hat{\xi}^{n-1,+}) \\
&\qquad\le \tfrac{1}{2}(\phi^{n-1}\tilde{\xi}^{n-1},\tilde{\xi}^{n-1}) + \tfrac{1}{2}(\phi^{n-1}\hat{\xi}^{n-1,+},\hat{\xi}^{n-1,+}) \\
&\qquad\le \tfrac{1}{2}(\phi^{n-1}\tilde{\xi}^{n-1},\tilde{\xi}^{n-1}) + \tfrac{1}{2}(\phi^n\xi^n,\xi^n) + Q\|\xi^n\|_0^2\,\Delta t^n,
\end{aligned}
\tag{6.25}
$$

where Q depends also on $\|\phi\|_{W^{1,\infty}(\Omega\times J)}$. Since $(\phi^n\xi^n,\xi^n) = (\phi^n\tilde{\xi}^n,\xi^n)$,

$$
\begin{aligned}
&(\phi^n\tilde{\xi}^n,\xi^n) - (\phi^{n-1}\tilde{\xi}^{n-1},\hat{\xi}^{n-1,+}) \\
&\quad\ge \tfrac{1}{2}\big[(\phi^n\tilde{\xi}^n,\xi^n) - (\phi^{n-1}\tilde{\xi}^{n-1},\tilde{\xi}^{n-1})\big] - Q\|\xi^n\|_0^2\,\Delta t^n \\
&\quad= \tfrac{1}{2}\big[(\phi^n\tilde{\xi}^n,\tilde{\xi}^n) - (\phi^{n-1}\tilde{\xi}^{n-1},\tilde{\xi}^{n-1})\big] - Q\|\xi^n\|_0^2\,\Delta t^n - \tfrac{1}{2}\big(\phi^n(\tilde{\xi}^n-\xi^n),\tilde{\xi}^n\big).
\end{aligned}
\tag{6.26}
$$

We have now extracted the collapsing part of our expression. The last term above must be controlled.

LEMMA 5. *For any element R,*

$$\big(\phi^n(\tilde{\xi}^n-\xi^n),\tilde{\xi}^n\big) = \big\|\sqrt{\phi^n}(\tilde{\xi}^n-\xi^n)\big\|_0^2 \tag{6.27}$$

$$\big\|\sqrt{\phi^n}\xi^n\big\|_0 \le \big\|\sqrt{\phi^n}\tilde{\xi}^n\big\|_0, \tag{6.28}$$

$$\|(D^n)^{1/2}\nabla\tilde{\xi}^n\|_{0,R} \le \|(D^n)^{-1/2}\zeta^n\|_{0,R}, \tag{6.29}$$

$$\big\|\sqrt{\phi^n}(\tilde{\xi}^n-\xi^n)\big\|_{0,R} \le K\|\nabla\tilde{\xi}^n\|_{0,R}\,h, \tag{6.30}$$

where K depends on the positive upper and lower bounds for D^n and ϕ^n.

Proof. By (3.2) and (4.10), we note that

$$\big(\phi^n(\tilde{\xi}^n-\xi^n),1\big)_R = 0, \tag{6.31a}$$

$$(D^n\nabla\tilde{\xi}^n + \zeta^n,\nabla\omega)_R = 0,\quad \omega\in\tilde{\mathcal{W}}_h. \tag{6.31b}$$

Since ξ^n is constant on R, (6.27)–(6.28) follow immediately from (6.31a). Taking $\omega = \tilde{\xi}^n$ in (6.31b), we obtain (6.29). Finally, for a good choice of constant W,

$$\begin{aligned}\big(\phi^n(\tilde{\xi}^n - \xi^n), \tilde{\xi}^n - \xi^n\big)_R &= \big(\phi^n(\tilde{\xi}^n - \xi^n), \tilde{\xi}^n - W\big)_R \\ &\le Q\big\|\sqrt{\phi^n}(\tilde{\xi}^n - \xi^n)\big\|_{0,R}\|\tilde{\xi}^n - W\|_{0,R} \\ &\le Q\big\|\sqrt{\phi^n}(\tilde{\xi}^n - \xi^n)\big\|_{0,R}\|\nabla\tilde{\xi}^n\|_{0,R}\, h,\end{aligned}$$

and (6.30) follows. □

We conclude that

$$\big|\big(\phi^n(\tilde{\xi}^n - \xi^n), \tilde{\xi}^n\big)\big| \le Q\|(D^n)^{-1/2}\zeta^n\|_0^2\, h^2, \tag{6.32}$$

where Q depends on the positive upper and lower bounds for D^n and ϕ^n. Later we will use our assumed relation between Δt and h to control this term.

It remains to analyze of the last two terms on the right-hand side of (6.24). First

$$\begin{aligned}-\,(\phi^n\tilde{\eta}^n, \xi^n) &= -(\phi^{n-1}\tilde{\eta}^{n-1}, \xi^n) - \int_{t^{n-1}}^{t^n} \big((\phi\tilde{\eta})_t, \xi^n\big)\, dt \\ &\le -(\phi^{n-1}\tilde{\eta}^{n-1}, \xi^n) + Q\Big\{\int_{t^{n-1}}^{t^n} \big[\|\tilde{\eta}\|_0^2 + \|\tilde{\eta}_t\|_0^2\big]\, dt + \|\xi^n\|_0^2\,\Delta t^n\Big\},\end{aligned} \tag{6.33}$$

so we are left with

$$\begin{aligned}&(\phi^{n-1}\tilde{\eta}^{n-1}, \hat{\xi}^{n-1,+} - \xi^n) \\ &\quad= \big((\check{\phi\tilde{\eta}})^{n-1,+}\mathcal{J}_n - \phi^{n-1}\tilde{\eta}^{n-1}, \xi^n\big) \\ &\quad\le \big((\check{\phi\tilde{\eta}})^{n-1,+} - \phi^{n-1}\tilde{\eta}^{n-1}, \xi^n\big) + Q\{\|\tilde{\eta}^{n-1}\|_0^2 + \|\xi^n\|_0^2\}\Delta t^n.\end{aligned} \tag{6.34}$$

The first term on the far right-hand side is the most difficult to estimate. In the standard analysis of the modified method of characteristics [18], one extracts a negative norm of the left-hand side of the inner product; however, in the characteristics-mixed method, ξ^n is discontinuous and therefore not in $H^1(\Omega)$. We begin by finding an $H^1(\Omega)$-function approximately equal to ξ^n.

LEMMA 6. *There is some $\Xi^n \in H^1(\Omega)$ and some constant K depending on upper and lower bounds for D and on $\|D\|_{L^\infty(J;W^{1,\infty}(\Omega))}$ but independent of h and n such that*

$$\|\Xi^n\|_1 \le K\{\|\xi^n\|_{-1} + \|\zeta^n\|_0\} \tag{6.35}$$

and, for sufficiently small h,

$$\|\Xi^n - \xi^n\|_0 \le K\{\|\xi^n\|_{-1} + \|\zeta^n\|_0\}\, h, \tag{6.36}$$

where $\|\cdot\|_{-1}$ denotes the norm of the dual to $H^1(\Omega)$.

Proof. Let $\mathcal{M}_h$ denote the standard Galerkin space of continuous, piecewise linear, bilinear, or trilinear functions (as appropriate) defined by its nodal values over the grid $\mathcal{T}_h$. Then define $\Xi^n \in \mathcal{M}_h$ by

$$(\Xi^n - \xi^n, \chi) + (D^n\nabla\Xi^n + \zeta^n, \nabla\chi) = 0, \quad \chi \in \mathcal{M}_h. \tag{6.37}$$

With $\chi = \Xi^n$, we obtain (6.35) (and therefore also existence and uniqueness of Ξ^n).

Let ψ solve the dual problem

$$\psi - \nabla\cdot(D^n\nabla\psi) = \Xi^n - \xi^n \quad \text{in } \Omega,$$
$$-D^n\nabla\psi\cdot\nu = 0 \quad \text{on } \partial\Omega,$$

for which

$$\|\psi\|_2 \le Q\|\Xi^n - \xi^n\|_0. \tag{6.38}$$

Then for any $\chi \in \mathcal{M}_h$,

$$\begin{aligned}\|\Xi^n - \xi^n\|_0^2 &= \big(\Xi^n - \xi^n, \psi - \nabla\cdot(D^n\nabla\psi)\big)\\ &= (\Xi^n - \xi^n, \psi) + (D^n\nabla\Xi^n, \nabla\psi) + \big(\xi^n, \nabla\cdot(D^n\nabla\psi)\big)\\ &= (\Xi^n - \xi^n, \psi - \chi) + \big(D^n\nabla\Xi^n, \nabla(\psi-\chi)\big)\\ &\quad + \big(\xi^n, \nabla\cdot(D^n\nabla\psi)\big) - (\zeta^n, \nabla\chi).\end{aligned} \tag{6.39}$$

Since $\xi^n \in \mathcal{W}_h$, using (6.1b),

$$\begin{aligned}\big(\xi^n, \nabla\cdot(D^n\nabla\psi)\big) &= \big(\xi^n, \nabla\cdot\pi_h(D^n\nabla\psi)\big) = \big((D^n)^{-1}\zeta^n, \pi_h(D^n\nabla\psi)\big)\\ &= (\zeta^n, \nabla\psi) + \big((D^n)^{-1}\zeta^n, \pi_h(D^n\nabla\psi) - D^n\nabla\psi\big).\end{aligned} \tag{6.40}$$

Combining this with (6.39) yields for a good choice of $\chi \in \mathcal{M}_h$ that

$$\begin{aligned}\|\Xi^n - \xi^n\|_0^2 &= (\Xi^n - \xi^n, \psi - \chi) + \big(D^n\nabla\Xi^n + \zeta^n, \nabla(\psi-\chi)\big)\\ &\quad + \big((D^n)^{-1}\zeta^n, \pi_h(D^n\nabla\psi) - D^n\nabla\psi\big)\\ &\le Q\big\{\|\Xi^n - \xi^n\|_0\, h + \|D^n\nabla\Xi^n\|_0 + \|\zeta^n\|_0\big\}\|\psi\|_2\, h,\end{aligned} \tag{6.41}$$

using (5.4). With (6.38) and (6.35), (6.36) follows for a sufficiently small h. □

Note that for any $\psi \in L^2(\Omega)$,

$$\begin{aligned}(\psi, \xi^n) &= (\psi, \Xi^n) + (\psi, \xi^n - \Xi^n)\\ &\le \|\psi\|_{-1}\|\Xi^n\|_1 + \|\psi\|_0\|\xi^n - \Xi^n\|_0\\ &\le Q\{\|\psi\|_{-1} + \|\psi\|_0\, h\}\{\|\xi^n\|_{-1} + \|\zeta^n\|_0\},\end{aligned}$$

where Q depends on $\|D\|_{L^\infty(J;W^{1,\infty}(\Omega))}$. Hence,

$$\begin{aligned}&\big(\check{(\phi\tilde\eta)}^{n-1,+} - \phi^{n-1}\tilde\eta^{n-1}, \xi^n\big)\\ &\quad\le Q\big\{\|\check{(\phi\tilde\eta)}^{n-1,+} - \phi^{n-1}\tilde\eta^{n-1}\|_{-1}\\ &\qquad + \|\check{(\phi\tilde\eta)}^{n-1,+} - \phi^{n-1}\tilde\eta^{n-1}\|_0\, h\big\}\{\|\xi^n\|_0 + \|\zeta^n\|_0\}.\end{aligned} \tag{6.42}$$

By [18, equations (4.32)–(4.37)] (cf. [13, equation (3.37)]),

$$\|\check{(\phi\tilde\eta)}^{n-1,+} - \phi^{n-1}\tilde\eta^{n-1}\|_{-1} \le Q\|\tilde\eta^{n-1}\|_0\, \Delta t^n, \tag{6.43}$$

and by Lemma 2,

$$\|\check{(\phi\tilde\eta)}^{n-1,+} - \phi^{n-1}\tilde\eta^{n-1}\|_0\, h \le Q\|\tilde\eta^{n-1}\|_0\, h. \tag{6.44}$$

This last estimate is not optimal (discontinuities in $\tilde\eta$ preclude the use of Lemma 3). We have that

$$\begin{aligned}&\big(\check{(\phi\tilde\eta)}^{n-1,+} - \phi^{n-1}\tilde\eta^{n-1}, \xi^n\big)\\ &\quad\le Q\big\{\big[\|\xi^n\|_0^2 + \|\zeta^n\|_0^2 + \|\tilde\eta^{n-1}\|_0^2\big]\,\Delta t^n + \|\tilde\eta^{n-1}\|_0^2\, h^2\,(\Delta t^n)^{-1}\big\}.\end{aligned} \tag{6.45}$$

Combining (6.24) with (6.26) and (6.32), and with (6.33)–(6.34) and (6.45), we obtain that

$$
\begin{aligned}
&\tfrac{1}{2}\big[(\phi^n\tilde{\xi}^n,\tilde{\xi}^n)-(\phi^{n-1}\tilde{\xi}^{n-1},\tilde{\xi}^{n-1})\big]\\
&\qquad+\big((D^n)^{-1}\zeta^n,\zeta^n\big)\,\big(\tfrac{1}{2}\Delta t^n-Q'h^2\big)+\int_{t^{n-1}}^{t^n}(f_2\hat{\xi},\hat{\xi})\,dt\\
&\quad\le Q\Big\{\Big[\|\tilde{\xi}^n\|_0^2+\|\eta^n\|_0^2\,(h^2+\Delta t^n)+\|\tilde{\eta}^n\|_0^2+\|\tilde{\eta}^{n-1}\|_0^2\\
&\qquad\quad+\int_{t^{n-1}}^{t^n}\big[\,\|c\|_1^2+\|c_\tau\|_0^2\,\big]\,dt\,\Delta t^n\\
&\qquad\quad+\int_{t^{n-1}}^{t^n}\big[\,\|(\nabla\cdot z)_\tau\|_0^2+\|\nabla\cdot z\|_0^2\,\big]\,dt\,\Delta t^n\Big]\Delta t^n\\
&\qquad+\int_{t^{n-1}}^{t^n}\big[\,\|\tilde{\eta}\|_0^2+\|\tilde{\eta}_t\|_0^2\,\big]\,dt+\|\tilde{\eta}^{n-1}\|_0^2\,h^2\,(\Delta t^n)^{-1}\Big\},
\end{aligned}
\tag{6.46}
$$

where Q' is the generic constant in (6.32). Since we have assumed that $\Delta t^n \geq K'h^{3/2}$,

$$Q'h^2 \le Q'K'\Delta t^n\,h^{1/2}$$

is negligible for h sufficiently small. Therefore, a summation on n, an application of Gronwall's inequality (provided that Δt^n is sufficiently small), and Theorem 2 yield that

$$
\begin{aligned}
&\max_n\|\tilde{\xi}^n\|_0^2+\sum_{n=1}^N\|\zeta^n\|_0^2\,\Delta t^n+\int_J(f_2\hat{\xi},\hat{\xi})\,dt\\
&\quad\le Q\Big\{\int_J\big[\,\|c\|_1^2+\|c_\tau\|_0^2+\|(\nabla\cdot z)_\tau\|_0^2+\|\nabla\cdot z\|_0^2\,\big]\,dt\,(\Delta t)^2\\
&\qquad+\max_{t\in J}\big[\,\|z\|_1^2+\|\nabla\cdot z\|_1^2\,\big]\,\big(h^4+h^2\,\Delta t+h^6\,(\Delta t)^{-2}\big)\\
&\qquad+\int_J\big[\,\|z_t\|_1^2+\|\nabla\cdot z_t\|_1^2\,\big]\,dt\,h^4+\|\tilde{\xi}^0\|_0^2\Big\}\\
&\quad\le Q(c)\{h^4+(\Delta t)^2+h^6\,(\Delta t)^{-2}+\|\tilde{\xi}^0\|_0^2\},
\end{aligned}
\tag{6.47}
$$

where $Q(c)$ has the form of $K(c)$. Since

$$h^6\,(\Delta t)^{-2}\le (K')^{-4}(\Delta t)^2,$$

the first half of Theorem 1 results.

If $u = 0$, then $(\check{\phi\tilde{\eta}})^{n-1,+} = \phi^{n-1}\tilde{\eta}^{n-1}$ and the term in (6.44) vanishes. With $\epsilon'\Delta t^n \geq h^2$, for ϵ' sufficiently small, $\epsilon'Q'$ is negligible compared to $\frac{1}{2}$, and we can obtain the last half of Theorem 1.

7. A stability result. As a corollary to the proof of Theorem 1, we have the following stability result independent of any assumptions on the true solution.

THEOREM 3. *Assume (A3)–(A6) and that $f_1 \in L^2(\Omega\times J)$. There is some constant $\epsilon' > 0$ such that if $\epsilon'\Delta t^n \geq h^2$, then for h and Δt sufficiently small,*

$$\max_n\{\|\tilde{c}_h^n\|_0+\|c_h^n\|_0\}+\Big\{\sum_{n=1}^N\|z_h^n\|_0^2\,\Delta t^n\Big\}^{1/2}\le K\{\|f_1\|_{L^2(\Omega\times J)}+\|\tilde{c}_h^0\|_0\},\tag{7.1}$$

for some constant K.

Proof. Set $W = c_h^n$ and $V = Z_h^n$ in (4.2) and combine with (3.2a) to obtain that

$$\begin{aligned}(7.2)\quad &(\phi^n \tilde{c}_h^n, c_h^n) - (\phi^{n-1}\tilde{c}_h^{n-1}, \hat{c}_h^{n-1,+}) + \big((D^n)^{-1} z_h^n, z_h^n\big)\Delta t^n + \int_{t^{n-1}}^{t^n} (f_2 \hat{c}_h, \hat{c}_h)\, dt \\ &= \int_{t^{n-1}}^{t^n} (f_1, \hat{c}_h)\, dt \le Q\Big\{ \int_{t^{n-1}}^{t^n} \|f_1\|_0^2\, dt + \|c_h^n\|_0^2\, \Delta t^n \Big\},\end{aligned}$$

using Lemma 2.

An argument analogous to (6.25)–(6.26) yields that

$$\begin{aligned}(7.3)\quad (\phi^{n-1}\tilde{c}_h^{n-1}, \hat{c}_h^{n-1,+}) &\le \tfrac{1}{2}(\phi^{n-1}\tilde{c}_h^{n-1}, \tilde{c}_h^{n-1}) + \tfrac{1}{2}(\phi^{n-1}\hat{c}_h^{n-1,+}, \hat{c}_h^{n-1,+}) \\ &\le \tfrac{1}{2}(\phi^{n-1}\tilde{c}_h^{n-1}, \tilde{c}_h^{n-1}) + \tfrac{1}{2}(\phi^n c_h^n, c_h^n) + Q\|c_h^n\|_0^2\, \Delta t^n,\end{aligned}$$

and so

$$\begin{aligned}(7.4)\quad &(\phi^n \tilde{c}_h^n, c_h^n) - (\phi^{n-1}\tilde{c}_h^{n-1}, \hat{c}_h^{n-1,+}) \\ &\ge \tfrac{1}{2}\big[(\phi^n \tilde{c}_h^n, c_h^n) - (\phi^{n-1}\tilde{c}_h^{n-1}, \tilde{c}_h^{n-1})\big] - Q\|c_h^n\|_0^2\, \Delta t^n \\ &= \tfrac{1}{2}\big[(\phi^n \tilde{c}_h^n, \tilde{c}_h^n) - (\phi^{n-1}\tilde{c}_h^{n-1}, \tilde{c}_h^{n-1})\big] - Q\|c_h^n\|_0^2\, \Delta t^n - \tfrac{1}{2}\big(\phi^n(\tilde{c}_h^n - c_h^n), \tilde{c}_h^n\big).\end{aligned}$$

Since (6.31) has the same form as (3.2), we obtain an analogue of Lemma 5 with c_h^n replacing ξ^n and z_h^n replacing ζ^n. Therefore,

$$(7.5)\quad \|c_h^n\|_0 \le Q\|\tilde{c}_h^n\|_0$$

and

$$(7.6)\quad \big|\big(\phi^n(\tilde{c}_h^n - c_h^n), \tilde{c}_h^n\big)\big| \le Q\|(D^n)^{-1/2} z_h^n\|_0^2\, h^2 \le \epsilon' Q\|(D^n)^{-1/2} z_h^n\|_0^2\, \Delta t^n.$$

The Theorem results from combining (7.2) with (7.4)–(7.6), taking ϵ' sufficiently small, and applying Gronwall. □

8. A remark on the use of other mixed spaces. We close the paper by remarking that the characteristics-mixed method can be defined for other mixed spaces. If $\mathcal{W}_h$ does not consist of piece-wise constants, then $\hat{W}$ must be interpreted carefully, since its spatial variation must be carried along the characteristics.

However, we are unable to modify our proof given above to obtain any better results on the convergence of the postprocessed concentration. Our proof suggests that the diffusive flux can be better approximated by higher order methods, but only to order $h^{3/2}$. It would be natural to use the lowest order spaces of Brezzi-Douglas-Marini [6] and Brezzi-Douglas-Durán-Fortin [5] to obtain a better approximation to z, though it is difficult to see why this would be desired. Since these higher order approximation spaces are computationally more difficult to implement, it is uncertain that their use is of any benefit.

REFERENCES

[1] T. Arbogast, A. Chilakapati, and M. F. Wheeler, *A characteristic-mixed method for contaminant transport and miscible displacement*, in Computational Methods in Water Resources IX, Vol. 1: Numerical Methods in Water Resources, T. F. Russell, R.E. Ewing, C. A. Brebbia, W. G. Gray, and G. F. Pindar, eds., Computational Mechanics Publications, Southampton, U.K., 1992, pp. 77-84.

[2] J. B. Bell, C. N. Dawson, and G. R. Shubin, *An unsplit higher-order Godunov scheme for scalar conservation laws in two dimensions*, J. Comp. Physics, 74 (1988), pp. 1–24.

[3] J. H. Bramble and S. R. Hilbert, *Estimation of linear functionals on Sobolev spaces with applications to Fourier transforms and spline interpolation*, SIAM J. Numer. Anal., 7 (1970), pp. 112–124.

[4] J. H. Bramble and S. R. Hilbert, *Bounds for a class of linear functionals with applications to Hermite interpolation*, Numer. Math., 15 (1971), pp. 362–369.

[5] F. Brezzi, J. Douglas, Jr., R. Durán, and M. Fortin, *Mixed finite elements for second order elliptic problems in three variables*, Numer. Math., 51 (1987), pp. 237–250.

[6] F. Brezzi, J. Douglas, Jr., and L. D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math., 47 (1985), pp. 217–235.

[7] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, Berlin, 1991.

[8] M. A. Celia, T. F. Russell, I. Herrera, and R. E. Ewing, *An Eulerian-Lagrangian localized adjoint method for the advection-diffusion equation*, Advances in Water Resources, 13 (1990), pp. 187–206.

[9] G. Chavent and J. Jaffré, *Mathematical models and finite elements for reservoir simulation*, Elsevier Science Publishers, New York, 1986.

[10] Zhangxin Chen and J. Douglas, Jr., *Prismatic mixed finite elements for second order elliptic problems*, Calcolo, 26 (1989), pp. 135–148.

[11] P. Colella and P. R. Woodward, *The piecewise-parabolic method (PPM) for gas-dynamical simulations*, J. Comp. Physics, 54 (1984), pp. 174–201.

[12] B. L. Darlow, R. E. Ewing, and M. F. Wheeler, *Mixed finite element methods for miscible displacement problems in porous media*, SPE 10501, Soc. Petrol. Eng. J., 24 (1984), pp. 391–398.

[13] C. N. Dawson, T. F. Russell, and M. F. Wheeler, *Some improved error estimates for the modified method of characteristics*, SIAM J. Numer. Anal., 26 (1989), pp. 1487–1512.

[14] J. Douglas, Jr., and T. Dupont, *Galerkin methods for parabolic equations*, SIAM J. Numer. Anal., 7 (1970), pp. 575–626.

[15] J. Douglas, Jr., and J. E. Roberts, *Global estimates for mixed methods for second order elliptic problems*, Math. Comp., 44 (1985), pp. 39–52.

[16] J. Douglas, Jr., and T. F. Russell, *Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures*, SIAM J. Numer. Anal., 19 (1982), pp. 871–885.

[17] T. Dupont and L. R. Scott, *Polynomial approximation of functions in Sobolev space*, Math. Comp., 34 (1980), pp. 441–463.

[18] R. E. Ewing, T. F. Russell, and M. F. Wheeler, *Convergence analysis of an approximation of miscible displacement in porous media by mixed finite elements and a modified method of characteristics*, Comp. Meth. Appl. Mech. Eng., 47 (1984), pp. 73–92.

[19] C. Johnson, *Streamline diffusion methods for problems in fluid mechanics*, in Finite Elements in Fluids VI, Wiley, New York, 1986.

[20] B. van Leer, *Towards the ultimate conservative difference scheme via second-order sequel to Godonov's methods*, J. Comp. Physics, 32 (1979), pp. 101–136.

[21] J. C. Nedelec, *Mixed finite elements in* $\mathbb{R}^3$, Numer. Math., 35 (1980), pp. 315–341.

[22] P-A. Raviart and J. M. Thomas, *A mixed finite element method for second order elliptic problems*, in Mathematical Aspects of the Finite Element Method, Lecture Notes in Mathematics 606, Springer-Verlag, Berlin, 1977, pp. 292–315.

[23] R. Stenberg, *Some new families of finite elements for the Stokes equations*, Numer. Math., 56 (1990), pp. 827–838.

[24] J. M. Thomas, *Sur l'analyse numérique des methodes d'éléments finis hybrides et mixtes*, in Thèse, Université Pierre-et-Marie Curie, Paris, 1977.

[25] M. R. Todd, P. M. O'Dell, and G. J. Hirasaki, *Methods for increased accuracy in numerical reservoir simulators*, Soc. Petrol. Eng. J., 12 (1972), pp. 515–530.

[26] M. F. Wheeler, *A priori* L_2 *error estimates for Galerkin approximations to parabolic partial differential equations*, SIAM J. Numer. Anal., 10 (1973), pp. 723–759.