CONVERGENCE ANALYSIS OF WAVELET SCHEMES FOR CONVECTION-REACTION EQUATIONS UNDER MINIMAL REGULARITY ASSUMPTIONS[*]

JIANGGUO LIU[†], BOJAN POPOV[‡], HONG WANG[§], AND RICHARD E. EWING[†]

Abstract. In this paper, we analyze convergence rates of wavelet schemes for time-dependent convection-reaction equations within the framework of the Eulerian–Lagrangian localized adjoint method (ELLAM). Under certain minimal assumptions that guarantee H^1-regularity of exact solutions, we show that a generic ELLAM scheme has a convergence rate $\mathcal{O}(h/\sqrt{\Delta t} + \Delta t)$ in L^2-norm. Then, applying the theory of operator interpolation, we obtain error estimates for initial data with even lower regularity. Namely, it is shown that the error of such a scheme is $\mathcal{O}((h/\sqrt{\Delta t})^\theta + (\Delta t)^\theta)$ for initial data in a Besov space $B_{2,q}^\theta(0 < \theta < 1, 0 < q \leq \infty)$. The error estimates are a priori and optimal in some cases. Numerical experiments using orthogonal wavelets are presented to illustrate the theoretical estimates.

Key words. characteristic method, convection-reaction equation, convergence rate, Eulerian–Lagrangian method, wavelet method

AMS subject classifications. 65M12, 65M25, 65M60, 76M25, 76S05

DOI. 10.1137/S0036142903433832

1. Introduction. This paper is concerned with convergence rates of the wavelet schemes established in [24] for an initial value problem (IVP) to the following multi-dimensional linear convection-reaction equation

$$(1.1) \qquad \begin{cases} u_t + \nabla \cdot (\mathbf{V}u) + Ru = f(\mathbf{x}, t), & (\mathbf{x}, t) \in \mathbb{R}^d \times (0, T], \\ u(\mathbf{x}, 0) = u_0(\mathbf{x}), & \mathbf{x} \in \mathbb{R}^d, \end{cases}$$

where $u(\mathbf{x}, t)$ is the unknown concentration function, $\mathbf{V}(\mathbf{x}, t)$ is a fluid velocity field, $R(\mathbf{x}, t)$ is a first order reaction coefficient, $f(\mathbf{x}, t)$ is a source/sink term, and $u_0(\mathbf{x})$ is a prescribed initial condition. It is assumed that $u_0(\mathbf{x})$ and $f(\mathbf{x}, t)$ are compactly supported, and hence so is the exact solution $u(\mathbf{x}, t)$ for any finite time.

Convection-dominated reactive transport equations arise from remediation of subsurface contamination, nuclear waste disposal, biodegradation, numerical simulation of petroleum reservoir, and many other applications. The solutions to these types of problems usually are not smooth and raise serious challenges to numerical methods. Standard finite difference or finite element methods (FEMs) produce either excessively oscillatory or smeared solutions. Therefore, many special schemes have been developed to overcome these difficulties.

Characteristics-based methods were developed in the late 1970s and the early 1980s to solve convection-dominated problems. Systematic study including error estimates for the FEM in the case of nondegenerate diffusion was given in [17]. Improved estimates for such problems were given in [12] with a special discussion for

degenerate diffusion. In the case of pure advection, [11] gives optimal rates of convergence for the piecewise linear FEM but with the restrictions of very small time steps $(\Delta t = \mathcal{O}(h^2))$ and sufficiently smooth exact solutions. There are many other results for convection-dominated diffusion problems, but in most cases the derived estimates adversely depend on the size of the diffusion parameter. In a recent paper [2], Bause and Knabner derived error estimates for Lagrangian–Galerkin methods, which are uniform with respect to the diffusion parameter, and therefore they remark that their results can carry over to the limit case of pure convection.

Among the existing characteristic methods, the Eulerian–Lagrangian localized adjoint method (ELLAM) [5] holds some advantages. It symmetrizes the governing equation, naturally incorporates boundary conditions, and conserves mass. However, ELLAM introduces further difficulties to the already fairly complicated analyses of characteristic methods. It was shown in [19, 25] that within the ELLAM framework, the piecewise linear FEM with uniform spatial partition has an optimal error estimate; in other words, it has a convergence rate $\mathcal{O}(h^2 + \Delta t)$, under the assumptions that $u \in L^\infty(0, T; H^3(\Omega))$ and $u_t \in L^2(0, T; H^2(\Omega))$. Many other papers also assume the exact solution is at least in H^2 in space. However, it is clear that the requirement $u \in L^\infty(0, T; H^s(\Omega))$ for any initial condition $u_0 \in H^s(\Omega)$ will imply that the fluid velocity is, roughly speaking, s times differentiable.

In this paper, requirements on regularity of the solution and the velocity field will be significantly reduced. We shall prove that under certain minimal assumptions that guarantee H^1-smoothness of the exact solution, ELLAM schemes including the orthogonal wavelet schemes satisfy

$$(1.2) \qquad \max_{0 \le n \le N} \|u(\mathbf{x}, t_n) - U_h^n(\mathbf{x})\|_{L^2(\mathbb{R}^d)} \le C(h/\sqrt{\Delta t} + \Delta t),$$

where $u(\mathbf{x}, t_n)$ is the exact solution at time t_n, and $U_h^n(\mathbf{x})$ is the corresponding numerical solution with spatial step size h and temporal step size Δt. The constant C depends only on the norms of the velocity, reaction, source, and initial data, but not on the norm of the exact solution $u(\mathbf{x}, t)$ itself.

Applying the theory of operator interpolation, we could obtain the error estimates

$$(1.3) \qquad \max_{0 \le n \le N} \|u(\mathbf{x}, t_n) - U_h^n(\mathbf{x})\|_{L^2(\mathbb{R}^d)} \le C \left[(h/\sqrt{\Delta t})^\theta + (\Delta t)^\theta \right]$$

for initial data u_0 in a Besov space $B_{2,q}^\theta (0 < \theta < 1, 0 < q \le \infty)$, where C could be additionally dependent on θ. This extends our results to a wide class of data, including discontinuous initial conditions, moving sharp fronts, or shocks. Generally speaking, Besov spaces provide subtler characterization on regularity of functions than Sobolev spaces do. Sometimes the exact order of approximation accuracy can be described only when Besov spaces are used, as illustrated by our Example 2 in section 8. Efforts in applying Besov spaces to other problems in numerical analysis can also be observed in literature, e.g., a recent work by Bacuta, Bramble, and Xu [1].

Errors in the wavelet schemes come from truncation, characteristic tracking, quadrature, and round-off. In this paper, we disregard quadrature rule error in the computations of wavelet coefficients. That is, we assume numerical integration is exact, following the common practices of most researchers [12, 17, 25]. The classic book [6] (see sections 4.1 and 4.4 of that work) has a full discussion on quadrature errors in the finite element method for elliptic problems. For time-dependent problems, some discussions can be found in [21].

The rest of this paper is organized as follows: Section 2 outlines the main ideas in the ELLAM weak formulation and establishes our numerical schemes, including the orthogonal wavelet schemes, within the ELLAM framework. Section 3 presents error estimates for the numerical schemes under some minimal regularity assumptions on the given data that guarantee H^1-stability of the exact solutions (Theorem 1). The proof of Theorem 1 is presented in section 4. In section 5, we state and prove the stability lemmas about the exact solutions used in section 4. Section 6 extends the results on convergence rates of the numerical schemes to initial data in Besov spaces. In section 7, we discuss optimality of our error estimates. Numerical experiments are presented in section 8 to illustrate the theoretical results. Finally, section 9 concludes the paper with some remarks.

2. ELLAM schemes. In this section, we establish the ELLAM weak formulation for problem (1.1). Based on this weak formulation, we derive a generic (abstract) ELLAM scheme. Then we present the wavelet-ELLAM schemes using orthogonal wavelets.

2.1. ELLAM formulation. Let $[0, T]$ be the time period, N a positive integer, $\Delta t := T/N$, and $t_n = n\Delta t$ $(n = 0, \ldots, N)$ be a uniform partition of $[0, T]$. To establish a weak formulation for (1.1), we choose test function $w(\mathbf{x}, t)$ in such a way that it vanishes outside the space-time strip $\mathbb{R}^d \times (t_{n-1}, t_n]$ and is discontinuous in time at time t_{n-1}. Then integration by parts gives us

$$
(2.1) \quad
\begin{aligned}
&\int_{\mathbb{R}^d} u(\mathbf{x}, t_n) w(\mathbf{x}, t_n) d\mathbf{x} - \int_{t_{n-1}}^{t_n} \int_{\mathbb{R}^d} (u(w_t + \mathbf{V} \cdot \nabla w - Rw))(\mathbf{x}, t) d\mathbf{x} dt \\
&= \int_{\mathbb{R}^d} u(\mathbf{x}, t_{n-1}) w(\mathbf{x}, t_{n-1}^+) d\mathbf{x} + \int_{t_{n-1}}^{t_n} \int_{\mathbb{R}^d} f(\mathbf{x}, t) w(\mathbf{x}, t) d\mathbf{x} dt,
\end{aligned}
$$

where $w(\mathbf{x}, t_{n-1}^+) := \lim_{t \to t_{n-1}^+} w(\mathbf{x}, t)$ takes into account the fact that $w(\mathbf{x}, t)$ is discontinuous in time at time t_{n-1}.

To cancel the second term on the left side of (2.1), we require the test function to satisfy the adjoint equation

$$
(2.2) \quad w_t + \mathbf{V} \cdot \nabla w - Rw = 0.
$$

Solving the above problem yields an explicit expression for the test function:

$$
(2.3) \quad w(\mathbf{y}(s; \mathbf{x}, t_n), s) = w(\mathbf{x}, t_n) e^{\int_{t_n}^s R(\mathbf{y}(r; \mathbf{x}, t_n), r) dr}, \quad s \in (t_{n-1}, t_n],
$$

where the characteristic $\mathbf{y}(s; \mathbf{x}, t_n)$ passing through $(\mathbf{x}, t_n)$ is determined by

$$
(2.4) \quad
\begin{cases}
\dfrac{d\mathbf{y}}{ds} = \mathbf{V}(\mathbf{y}, s), \\
\mathbf{y}(s; \mathbf{x}, t_n)|_{s=t_n} = \mathbf{x}.
\end{cases}
$$

Then we are led to the reference equation

$$
(2.5) \quad
\begin{aligned}
\int_{\mathbb{R}^d} u(\mathbf{x}, t_n) w(\mathbf{x}, t_n) d\mathbf{x} &= \int_{\mathbb{R}^d} u(\mathbf{x}, t_{n-1}) w(\mathbf{x}, t_{n-1}^+) d\mathbf{x} \\
&+ \int_{t_{n-1}}^{t_n} \int_{\mathbb{R}^d} f(\mathbf{x}, t) w(\mathbf{x}, t) d\mathbf{x} dt.
\end{aligned}
$$

Let $\mathbf{x}^* = \mathbf{y}(t_{n-1}; \mathbf{x}, t_n)$. Applying (2.3), we can rewrite the first term on the right side of (2.5) as

$$
\begin{aligned}
(2.6) \qquad & \int_{\mathbb{R}^d} u(\mathbf{x}, t_{n-1}) w(\mathbf{x}, t_{n-1}^+) d\mathbf{x} \\
& = \int_{\mathbb{R}^d} u(\mathbf{x}^*, t_{n-1}) w(\mathbf{x}, t_n) e^{\int_{t_n}^{t_{n-1}} R(\mathbf{y}(s;\mathbf{x},t_n),s)ds} \mathbf{J}(\mathbf{x}^*, \mathbf{x}) d\mathbf{x},
\end{aligned}
$$

where $\mathbf{J}(\mathbf{x}^*, \mathbf{x})$ is the Jacobian of $\mathbf{x}^*$ with respect to $\mathbf{x}$.

2.2. A generic ELLAM scheme. Based on (2.3), we first approximate the test function $w(\mathbf{y}(s; \mathbf{x}, t_n), s)$ in the space-time strip $\mathbb{R}^d \times (t_{n-1}, t_n]$ by

$$
w(\mathbf{x}, t_n) \, e^{\int_{t_n}^s R(\mathbf{x},t_n)dr} \equiv w(\mathbf{x}, t_n) \, e^{R(\mathbf{x},t_n)(s-t_n)},
$$

and then use it to approximate the second (source) term on the right side of (2.5),

$$
\begin{aligned}
(2.7) \qquad & \int_{t_{n-1}}^{t_n} \int_{\mathbb{R}^d} f(\mathbf{y}, s) w(\mathbf{y}, s) d\mathbf{y} ds \\
& = \int_{\mathbb{R}^d} \int_{t_{n-1}}^{t_n} f(\mathbf{y}, s) w(\mathbf{y}, s) \mathbf{J}(\mathbf{y}, \mathbf{x}) ds d\mathbf{x} \\
& = \int_{\mathbb{R}^d} \int_{t_{n-1}}^{t_n} f(\mathbf{x}, t_n) w(\mathbf{x}, t_n) e^{R(\mathbf{x},t_n)(s-t_n)} ds d\mathbf{x} + E(f, w) \\
& = \int_{\mathbb{R}^d} f(\mathbf{x}, t_n) w(\mathbf{x}, t_n) G_n(\mathbf{x}) d\mathbf{x} + E(f, w),
\end{aligned}
$$

where $\mathbf{J}(\mathbf{y}, \mathbf{x})$ is the Jacobian of $\mathbf{y}$ with respect to $\mathbf{x}$,

$$
(2.8) \qquad G_n(\mathbf{x}) := \int_{t_{n-1}}^{t_n} e^{R(\mathbf{x},t_n)(s-t_n)} ds = \begin{cases} \dfrac{1 - e^{-R(\mathbf{x},t_n)\Delta t}}{R(\mathbf{x}, t_n)} & \text{if } R(\mathbf{x}, t_n) \neq 0, \\ \Delta t & \text{otherwise,} \end{cases}
$$

and $E(f, w)$ is the error term

$$
(2.9) \qquad E(f, w) := \int_{t_{n-1}}^{t_n} \int_{\mathbb{R}^d} f(\mathbf{y}, s) w(\mathbf{y}, s) d\mathbf{y} ds - \int_{\mathbb{R}^d} f(\mathbf{x}, t_n) w(\mathbf{x}, t_n) G_n(\mathbf{x}) d\mathbf{x}.
$$

In practice, exact tracking of characteristics is usually unavailable, and we have to resort to numerical means. All commonly used numerical methods, e.g., Euler and Runge–Kutta methods, can be employed to solve (2.4). Let $(\mathbf{x}^{**}, t_{n-1})$ be the numerical back-tracking image of $(\mathbf{x}, t_n)$; then $w(\mathbf{x}^{**}, t_{n-1}^+)$ can be approximated by $w(\mathbf{x}, t_n)e^{-R(\mathbf{x},t_n)\Delta t}$.

Now a generic ELLAM scheme can be established as follows: Let $V_h \subset L^2(\mathbb{R}^d)$ be an approximation subspace. Find $U_h^n(\mathbf{x}) \in V_h$ such that for any test function w with $w(\mathbf{x}, t_n) \in V_h$,

$$
\begin{aligned}
(2.10) \qquad & \int_{\mathbb{R}^d} U_h^n(\mathbf{x}) w(\mathbf{x}, t_n) d\mathbf{x} = \int_{\mathbb{R}^d} U_h^{n-1}(\mathbf{x}^{**}) w(\mathbf{x}, t_n) e^{-R(\mathbf{x},t_n)\Delta t} \mathbf{J}(\mathbf{x}^{**}, \mathbf{x}) d\mathbf{x} \\
& \qquad + \int_{\mathbb{R}^d} f(\mathbf{x}, t_n) w(\mathbf{x}, t_n) G_n(\mathbf{x}) d\mathbf{x},
\end{aligned}
$$

where $U_h^0(\mathbf{x})$ is an approximation to $u_0(\mathbf{x})$ from V_h obtained by some means. For example, it can be taken as the L^2-orthogonal projection of u_0 into V_h.

There are many choices for the approximation subspace V_h. It could be constructed through finite elements or wavelets.

2.3. Wavelet schemes. Let

$$\cdots \subset \mathcal{V}_{-1} \subset \mathcal{V}_0 \subset \mathcal{V}_1 \subset \cdots \subset \mathcal{V}_j \subset \mathcal{V}_{j+1} \subset \cdots \subset L^2(\mathbb{R})$$

be a multiresolution analysis in $L^2(\mathbb{R})$ generated by an orthogonal scaling function $\phi(x)$ and let $\psi(x)$ be the associated orthogonal wavelet [10]. We construct d-dimensional scaling function and wavelets through tensor products. Let $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d, \mathbf{k} = (k_1, \ldots, k_d) \in \mathbb{Z}^d, \mathbf{e} = (e_1, \ldots, e_d) \in \{0, 1\}^d =: \hat{E}$, and $E := \hat{E} \setminus \{\mathbf{0}\}$. Then we define scaling function $\mathbf{\Phi}_{j,\mathbf{k}}(\mathbf{x}) := \prod_{i=1}^d \phi_{j,k_i}(x_i)$ and the wavelets

$$\Psi_{j,\mathbf{k}}^{\mathbf{e}}(\mathbf{x}) := \prod_{i=1}^d \left(\phi_{j,k_i}(x_i)\right)^{1-e_i} \left(\psi_{j,k_i}(x_i)\right)^{e_i}, \qquad \mathbf{e} \in E.$$

Furthermore, we define $V_j := \bigotimes_{i=1}^d \mathcal{V}_j$ as the closed linear span in $L^2(\mathbb{R}^d)$ of all functions of the form $f_1(x_1) \cdots f_d(x_d)$, where $f_i \in \mathcal{V}_j$ for $i = 1, \ldots, d$. Then $\langle V_j \rangle_{j \in \mathbb{Z}}$ forms a multiresolution analysis in $L^2(\mathbb{R}^d)$.

Clearly, we can use any subspace V_j in the above multiresolution analysis as an approximation subspace V_h in the ELLAM scheme (2.10). This gives us an orthogonal wavelet-ELLAM scheme.

For numerical implementations, we assume $\Omega \subset \mathbb{R}^d$ to be a rectangular domain such that the support of the solution $u(\mathbf{x}, t)$ to problem (1.1) is contained in Ω for all $t \in [0, T]$. Let $J_c < J_f$ be the chosen coarsest and finest resolution levels. For all j with $J_c \leq j \leq J_f$, we define

$$(2.11) \qquad \Lambda_j := \{\mathbf{k} : \mathrm{supp}\mathbf{\Phi}_{j,\mathbf{k}} \cap \Omega \neq \emptyset\},$$

$$(2.12) \qquad \Lambda_{j,\mathbf{e}} := \{\mathbf{k} : \mathrm{supp}\Psi_{j,\mathbf{k}}^{\mathbf{e}} \cap \Omega \neq \emptyset\},$$

$$(2.13) \qquad \mathcal{S}_j(\Omega) := \mathrm{Span}\{\mathbf{\Phi}_{j,\mathbf{k}}(\mathbf{x}) \mid \mathbf{k} \in \Lambda_j\}.$$

Then we set $V_h = \mathcal{S}_{J_f}(\Omega)$ with $h = 1/2^{J_f}$. There are two equivalent choices for the basis functions of V_h:

- only the scaling functions at the finest level J_f;
- the scaling functions at the coarsest level J_c plus all wavelets from level J_c to level $J_f - 1$.

When only the scaling functions at the finest level are used as the basis functions, we are seeking $U_h^n(\mathbf{x}) \in V_h$ with

$$(2.14) \qquad U_h^n(\mathbf{x}) = \sum_{\mathbf{k} \in \Lambda_{J_f}} c_{J_f,\mathbf{k}}^n \mathbf{\Phi}_{J_f,\mathbf{k}}(\mathbf{x}).$$

The orthogonality of scaling functions implies that we have an explicit scheme and the coefficients are given by

$$(2.15) \qquad \begin{aligned} c_{J_f,\mathbf{k}}^n &= \int_\Omega U_h^{n-1}(\mathbf{x}^{**}) \mathbf{\Phi}_{J_f,\mathbf{k}}(\mathbf{x}) e^{-R(\mathbf{x},t_n)\Delta t} \mathbf{J}(\mathbf{x}^{**}, \mathbf{x}) d\mathbf{x} \\ &\quad + \int_\Omega f(\mathbf{x}, t_n) \mathbf{\Phi}_{J_f,\mathbf{k}}(\mathbf{x}) G_n(\mathbf{x}) d\mathbf{x}, \qquad \mathbf{k} \in \Lambda_{J_f}. \end{aligned}$$

This is the Scheme I discussed in [24].

When both the scaling functions at the coarsest level and the wavelets at fine levels are used as the basis functions, we get Scheme II with

$$(2.16) \qquad U_h^n(\mathbf{x}) = \sum_{\mathbf{k} \in \Lambda_{J_c}} c_{J_c,\mathbf{k}}^n \Phi_{J_c,\mathbf{k}}(\mathbf{x}) + \sum_{j=J_c}^{J_f-1} \sum_{\mathbf{k} \in \Lambda_{j,\mathbf{e}}} d_{j,\mathbf{k}}^{n,\mathbf{e}} \Psi_{j,\mathbf{k}}^{\mathbf{e}}(\mathbf{x}).$$

Scheme II is also an explicit scheme. By choosing $w(\mathbf{x}, t_n) = \Phi_{J_c,\mathbf{k}}(\mathbf{x})$ or $\Psi_{j,\mathbf{k}}^{\mathbf{e}}(\mathbf{x})$, respectively, we obtain, again through the orthogonality,

$$
\begin{aligned}
c_{J_c,\mathbf{k}}^n &= \int_\Omega U_h^{n-1}(\mathbf{x}^{**}) \Phi_{J_c,\mathbf{k}}(\mathbf{x}) e^{-R(\mathbf{x},t_n)\Delta t} \mathbf{J}(\mathbf{x}^{**}, \mathbf{x}) d\mathbf{x} \\
&\quad + \int_\Omega f(\mathbf{x}, t_n) \Phi_{J_c,\mathbf{k}}(\mathbf{x}) G_n(\mathbf{x}) d\mathbf{x}, \qquad \mathbf{k} \in \Lambda_{J_c}, \\
d_{j,\mathbf{k}}^{n,\mathbf{e}} &= \int_\Omega U_h^{n-1}(\mathbf{x}^{**}) \Psi_{j,\mathbf{k}}^{\mathbf{e}}(\mathbf{x}) e^{-R(\mathbf{x},t_n)\Delta t} \mathbf{J}(\mathbf{x}^{**}, \mathbf{x}) d\mathbf{x} \\
&\quad + \int_\Omega f(\mathbf{x}, t_n) \Psi_{j,\mathbf{k}}^{\mathbf{e}}(\mathbf{x}) G_n(\mathbf{x}) d\mathbf{x}, \qquad \mathbf{k} \in \Lambda_{j,\mathbf{e}}, \ \ J_c \le j \le J_f - 1.
\end{aligned}
$$

(2.17)

In Scheme II, the first part on the right side of (2.16) provides a basic approximation. As more fine terms in the second part come in, we obtain better approximations.

As we know, solutions to convection-dominated transport equations often admit steep fronts and even jump discontinuities within very small regions but are smooth outside these regions. On the other hand, one prominent feature of wavelets is their localization capability. The terms in the wavelet expansion with noticeable coefficients correspond to the rough regions of the solution around those local singularities. We can drop the terms with small coefficients that correspond to the smooth regions of the solution. Therefore, the number of unknowns to be solved will be reduced. In other words, an adaptive multilevel scheme with mass-conservative compression can be constructed. This is the Scheme III presented in [24], which is, in some sense, equivalent to the traditional FEM with local refinement.

Due to compression, the wavelet basis functions (or elements) used in Scheme III vary at different time steps but are adapted to the solution we are looking for. This is a typical case of nonlinear approximation [14]. Of course, Schemes I and II are still in the category of linear approximation and are the main target of this paper. Theoretical analysis on convergence rates of Scheme III is much harder and will be addressed in our future work.

As proven in [24], all three wavelet schemes are explicit and unconditionally stable. In other words, they are not subject to the severe restriction of the Courant–Friedrichs–Lewy (CFL) condition. This allows us to use relatively large time steps.

3. Error estimate for solutions with H^1-regularity. In this section, we derive an error estimate for the generic ELLAM scheme for exact solutions with only H^1-regularity.

Throughout this paper, we use L^p $(1 \le p \le \infty)$ to denote the standard Lebesgue spaces. Accordingly, W_p^k are the standard Sobolev spaces. When $p = 2$, we use H^k for W_2^k. For $1 \le q \le \infty$, we define

$$L^q(a, b; W_p^k) := \{u(\mathbf{x}, t) | u(\cdot, t) : (a, b) \mapsto W_p^k, \|u(\cdot, t)\|_{W_p^k} \in L^q(a, b)\}.$$

In addition, $(W_\infty^1(\mathbb{R}^d \times [0, T]))^d$ is the space of vector-valued functions whose components are in the space $W_\infty^1(\mathbb{R}^d \times [0, T])$.

Assumption A. The approximation subspace V_h used in the generic ELLAM scheme (2.10) has the following approximation property:

$$(3.1) \qquad \|u - P_h u\|_{L^2} \leq C_0 h \|u\|_{H^1} \qquad \forall u \in H^1(\mathbb{R}^d),$$

where C_0 is a constant independent of h and u. The above inequality is also called a Jackson-type inequality in literature.

Remark A. The above Jackson-type inequality is satisfied by commonly used wavelets, e.g., Daubechies' wavelets; see [14].

Assumption B.
  (i) $\mathbf{V} \in (W_\infty^1(\mathbb{R}^d \times [0, T]))^d$.
  (ii) $\mathrm{div}\mathbf{V} \in W_\infty^1(\mathbb{R}^d \times [0, T])$.
  (iii) $R \in W_\infty^1(\mathbb{R}^d \times [0, T])$.

Now we state our first theorem on the error estimate for exact solutions with only H^1-regularity.

THEOREM 1. *Let $u(\mathbf{x}, t)$ be the exact solution of problem (1.1) and let $U_h^n(\mathbf{x})$ be the numerical solution generated by the generic ELLAM scheme (2.10). Then under Assumptions A and B, the following error estimate in L^2-norm holds:*

$$\max_{0 \leq n \leq N} \|u(\mathbf{x}, t_n) - U_h^n(\mathbf{x})\|_{L^2(\mathbb{R}^d)}$$

$$(3.2) \qquad \leq C \Bigg[\frac{h}{\sqrt{\Delta t}} \left(\|u_0\|_{H^1(\mathbb{R}^d)} + \|f\|_{L^2(0,T;H^1(\mathbb{R}^d))} \right)$$

$$+ \Delta t \left(\|u_0\|_{L^2(\mathbb{R}^d)} + \|f\|_{L^2(0,T;L^2(\mathbb{R}^d))} + \|f_\tau\|_{L^2(0,T;L^2(\mathbb{R}^d))} \right) \Bigg],$$

where Δt and h are the temporal and spatial step sizes, respectively, and f_τ is the total derivative, i.e., the derivative along the characteristic direction. The constant C depends only on the final time T and the norms of $\mathbf{V}, \mathrm{div}\mathbf{V}, R$ in the corresponding spaces in Assumption B. When orthogonal wavelet schemes are used, $h = 1/2^J$ with J being the finest spatial resolution used in the wavelet schemes.

4. Proof of Theorem 1. In this section, we first estimate the error $E_n(w)$ defined in (4.5). Then we estimate the error involving the source term defined in (2.9). Applying a discrete Gronwall inequality, we derive the final error estimate stated in Theorem 1.

Let $u_n(\mathbf{x}) := u(\mathbf{x}, t_n)$ and let $P_h u_n$ be the L^2-orthogonal projection of u_n into the subspace V_h, that is,

$$(4.1) \qquad (P_h u_n, v_h) = (u_n, v_h) \qquad \forall \, v_h \in V_h.$$

Now with $P_h u_n, U_h^n \in V_h$, we have the orthogonality

$$(4.2) \qquad u_n - U_h^n = (u_n - P_h u_n) \oplus (P_h u_n - U_h^n),$$

$$(4.3) \qquad \|u_n - U_h^n\|_{L^2(\mathbb{R}^d)}^2 = \|u_n - P_h u_n\|_{L^2(\mathbb{R}^d)}^2 + \|P_h u_n - U_h^n\|_{L^2(\mathbb{R}^d)}^2.$$

Here $u_n - P_h u_n$ is the approximation error, which is completely determined by u_n and the chosen approximation subspace V_h. It is independent of the numerical scheme being used: a wavelet method or a traditional finite element method. Therefore we only need to bound the second term on the right side of (4.3).

Let w be a test function such that $w(\mathbf{x}, t_n) \in V_h$. Subtracting (2.10) from (2.5) and then applying (2.6) and (2.9), we obtain

$$(4.4) \qquad (P_h u_n - U_h^n, w(\mathbf{x}, t_n)) = E_n(w) + E(f, w),$$

where

$$
\begin{aligned}
(4.5) \qquad E_n(w) := & \int_{\mathbb{R}^d} u(\mathbf{x}^*, t_{n-1}) w(\mathbf{x}, t_n) e^{\int_{t_n}^{t_{n-1}} R(\mathbf{y}, s) ds} \mathbf{J}(\mathbf{x}^*, \mathbf{x}) d\mathbf{x} \\
& - \int_{\mathbb{R}^d} U_h^{n-1}(\mathbf{x}^{**}) w(\mathbf{x}, t_n) e^{-R(\mathbf{x}, t_n) \Delta t} \mathbf{J}(\mathbf{x}^{**}, \mathbf{x}) d\mathbf{x}.
\end{aligned}
$$

From now on, we shall use C to denote a constant that is independent of the spatial and temporal mesh sizes but may depend on the final time T and the norms of the velocity field and the reaction in the corresponding spaces in Assumption B, whereas C_0 will be used for an absolute constant that does not depend on any of the aforementioned terms. All these constants may take different values in different occurrences.

The following basic estimates are easy to verify and shall be repeatedly used in this section. When Δt is small enough, we have

(i) $\mathbf{J}(\mathbf{x}^*, \mathbf{x}) \leq 1 + C\Delta t$,

(ii) $\mathbf{J}(\mathbf{x}^{**}, \mathbf{x}) \leq 1 + C\Delta t$,

(iii) $e^{\int_{t_n}^{t_{n-1}} R(\mathbf{y}, s) ds} \leq 1 + C\Delta t$,

(iv) $e^{-R(\mathbf{x}, t_n) \Delta t} \leq 1 + C\Delta t$.

Recall that Assumption B, part (i), implies that the velocity field is Lipschitz in both space and time. When the Euler method is used for tracking characteristics, we have

$$(4.6) \qquad \mathbf{x}^{**} = \mathbf{x} + \mathbf{V}(\mathbf{x}, t_n)(t_{n-1} - t_n),$$

and hence

$$(4.7) \qquad |\mathbf{x}^* - \mathbf{x}^{**}| \leq C(\Delta t)^2,$$

where $C = C_0 \|\mathbf{V}\|_{(W_\infty^1(\mathbb{R}^d \times [0,T]))^d}$. Furthermore, the following estimate holds under Assumption B, parts (i) and (ii):

$$(4.8) \qquad |\mathbf{J}(\mathbf{x}^*, \mathbf{x}) - \mathbf{J}(\mathbf{x}^{**}, \mathbf{x})| \leq C(\Delta t)^2.$$

4.1. Estimate on error $E_n(w)$. For the error $E_n(w)$, we split it into 4 terms:

$$(4.9) \qquad E_n(w) = I_n^{(1)} + I_n^{(2)} + I_n^{(3)} + I_n^{(4)},$$

where

$$
\begin{aligned}
I_n^{(1)} := & \int_{\mathbb{R}^d} \left[u(\mathbf{x}^*, t_{n-1}) - u(\mathbf{x}^{**}, t_{n-1}) \right] w(\mathbf{x}, t_n) e^{\int_{t_n}^{t_{n-1}} R(\mathbf{y}, s) ds} \mathbf{J}(\mathbf{x}^*, \mathbf{x}) d\mathbf{x}, \\
I_n^{(2)} := & \int_{\mathbb{R}^d} u(\mathbf{x}^{**}, t_{n-1}) w(\mathbf{x}, t_n) e^{\int_{t_n}^{t_{n-1}} R(\mathbf{y}, s) ds} \left[\mathbf{J}(\mathbf{x}^*, \mathbf{x}) - \mathbf{J}(\mathbf{x}^{**}, \mathbf{x}) \right] d\mathbf{x}, \\
(4.10) \qquad I_n^{(3)} := & \int_{\mathbb{R}^d} u(\mathbf{x}^{**}, t_{n-1}) w(\mathbf{x}, t_n) \left[e^{\int_{t_n}^{t_{n-1}} R(\mathbf{y}, s) ds} - e^{-R(\mathbf{x}, t_n) \Delta t} \right] \mathbf{J}(\mathbf{x}^{**}, \mathbf{x}) d\mathbf{x}, \\
I_n^{(4)} := & \int_{\mathbb{R}^d} \left[u(\mathbf{x}^{**}, t_{n-1}) - U_h^{n-1}(\mathbf{x}^{**}) \right] w(\mathbf{x}, t_n) e^{-R(\mathbf{x}, t_n) \Delta t} \mathbf{J}(\mathbf{x}^{**}, \mathbf{x}) d\mathbf{x}.
\end{aligned}
$$

Note that $I_n^{(1)}$ and $I_n^{(2)}$ in (4.10) reflect the error from inexact tracking of characteristics. These two terms will vanish when exact tracking of characteristics is available.

Estimate on $I_n^{(1)}$ in (4.10). The H^1-stability of the exact solution (Lemma 4 in section 5) and estimate (4.7) yield

$$\left(\int_{\mathbb{R}^d} |u(\mathbf{x}^*, t_n) - u(\mathbf{x}^{**}, t_n)|^2 d\mathbf{x} \right)^{1/2} \leq C\|u_n\|_{H^1(\mathbb{R}^d)}(\Delta t)^2.$$

Therefore,

(4.11) $$|I_n^{(1)}| \leq C(\Delta t)^2 \|u_n\|_{H^1(\mathbb{R}^d)} \|w(\mathbf{x}, t_n)\|_{L^2(\mathbb{R}^d)}.$$

Estimate on $I_n^{(2)}$ in (4.10). Based on estimate (4.8), we have

(4.12) $$|I_n^{(2)}| \leq C(\Delta t)^2 \|u_{n-1}\|_{L^2(\mathbb{R}^d)} \|w(\mathbf{x}, t_n)\|_{L^2(\mathbb{R}^d)}.$$

Estimate on $I_n^{(3)}$ in (4.10). Let R_τ be the total derivative of R (along characteristic) and $\mathbf{z} = \mathbf{y}(r; \mathbf{x}, t_n)$ for $r \in [s, t_n]$; then

$$\left| e^{\int_{t_n}^{t_{n-1}} R(\mathbf{y}, s)ds} - e^{-R(\mathbf{x}, t_n)\Delta t} \right| \leq 2 \int_{t_{n-1}}^{t_n} \int_s^{t_n} |R_\tau(\mathbf{z}, r)| dr ds \leq \|R_\tau\|_\infty (\Delta t)^2.$$

But $R_\tau = \nabla R \cdot \mathbf{V} + R_t$, so we have

(4.13) $$|I_n^{(3)}| \leq C(\Delta t)^2 \|u_{n-1}\|_{L^2(\mathbb{R}^d)} \|w(\mathbf{x}, t_n)\|_{L^2(\mathbb{R}^d)}.$$

Estimate on $I_n^{(4)}$ in (4.10). It is easy to derive the following estimate:

(4.14) $$|I_n^{(4)}| \leq (1 + C\Delta t)\|u_{n-1} - U_h^{n-1}\|_{L^2(\mathbb{R}^d)} \|w(\mathbf{x}, t_n)\|_{L^2(\mathbb{R}^d)}.$$

Substitution of estimates (4.11), (4.12), (4.13), and (4.14) into (4.9) gives us an estimate on the error $E_n(w)$ defined in (4.5):

(4.15) $$|E_n(w)| \leq \left(\frac{1}{2} + C\Delta t \right) \left(\|u_{n-1} - U_h^{n-1}\|_{L^2(\mathbb{R}^d)}^2 + \|w(\mathbf{x}, t_n)\|_{L^2(\mathbb{R}^d)}^2 \right)$$
$$+ C(\Delta t)^3 \|u_{n-1}\|_{L^2(\mathbb{R}^d)}^2 + C\Delta t \|w(\mathbf{x}, t_n)\|_{L^2(\mathbb{R}^d)}^2,$$

where C depends on the final time T and the norms of $\mathbf{V}, \mathrm{div}\mathbf{V}, R$ in the corresponding spaces in Assumption B.

4.2. Estimate on source term. Now we estimate the error in the approximation to the source term. Let $\mathbf{y} = \mathbf{y}(s; \mathbf{x}, t_n), s \in [t_{n-1}, t_n]$, and $\mathbf{z} = \mathbf{y}(r; \mathbf{x}, t_n), r \in [s, t_n]$. According to (2.3), (2.7), and (2.9), we have

(4.16) $$E(f, w) = I_f^{(1)} + I_f^{(2)} + I_f^{(3)},$$

where

$$I_f^{(1)} := \int_{t_{n-1}}^{t_n} \int_{\mathbb{R}^d} f(\mathbf{y}, s) e^{\int_{t_n}^s R(\mathbf{z}, r)dr} w(\mathbf{x}, t_n)[\mathbf{J}(\mathbf{y}, \mathbf{x}) - 1] d\mathbf{x} ds,$$

(4.17) $$I_f^{(2)} := \int_{t_{n-1}}^{t_n} \int_{\mathbb{R}^d} f(\mathbf{y}, s) \left[e^{\int_{t_n}^s R(\mathbf{z}, r)dr} - e^{R(\mathbf{x}, t_n)(s - t_n)} \right] w(\mathbf{x}, t_n) d\mathbf{x} ds,$$

$$I_f^{(3)} := \int_{t_{n-1}}^{t_n} \int_{\mathbb{R}^d} [f(\mathbf{y}, s) - f(\mathbf{x}, t_n)] e^{R(\mathbf{x}, t_n)(s - t_n)} w(\mathbf{x}, t_n) d\mathbf{x} ds.$$

For convenience, let $J_n := [t_{n-1}, t_n]$.

Estimate on $I_f^{(1)}$ in (4.17). Applying the Cauchy–Schwarz inequality first in space and then in time gives

$$\left| \int_{t_{n-1}}^{t_n} \int_{\mathbb{R}^d} f(\mathbf{y}, s) e^{\int_{t_n}^{s} R(\mathbf{z}, r) dr} w(\mathbf{x}, t_n) [\mathbf{J}(\mathbf{y}, \mathbf{x}) - 1] d\mathbf{x} ds \right|$$

$$\leq \int_{t_{n-1}}^{t_n} 4\sqrt{2} \|\mathrm{div}\mathbf{V}\|_{\infty} (t_n - s) \|f(\cdot, s)\|_{L^2(\mathbb{R}^d)} \|w(\mathbf{x}, t_n)\|_{L^2(\mathbb{R}^d)} ds$$

$$\leq \frac{4\sqrt{2}}{\sqrt{3}} \|\mathrm{div}\mathbf{V}\|_{\infty} (\Delta t)^{\frac{3}{2}} \|f\|_{L^2(J_n; L^2(\mathbb{R}^d))} \|w(\mathbf{x}, t_n)\|_{L^2(\mathbb{R}^d)}.$$

Therefore, we have

$$(4.18) \qquad |I_f^{(1)}| \leq \frac{2\sqrt{2}}{\sqrt{3}} \|\mathrm{div}\mathbf{V}\|_{\infty} \left[(\Delta t)^2 \|f\|_{L^2(J_n, L^2(\mathbb{R}^d))}^2 + \Delta t \|w(\mathbf{x}, t_n)\|_{L^2(\mathbb{R}^d)}^2 \right].$$

Estimate on $I_f^{(2)}$ in (4.17). Similar to the above, we have

$$(4.19) \qquad |I_f^{(2)}| \leq \frac{2\sqrt{2}}{\sqrt{3}} \|R\|_{\infty} \left[(\Delta t)^2 \|f\|_{L^2(J_n; L^2(\mathbb{R}^d))}^2 + \Delta t \|w(\mathbf{x}, t_n)\|_{L^2(\mathbb{R}^d)}^2 \right].$$

Estimate on $I_f^{(3)}$ in (4.17). Let f_τ be the total derivative of f; then

$$\left| \int_{\mathbb{R}^d} [f(\mathbf{y}, s) - f(\mathbf{x}, t_n)] e^{R(\mathbf{x}, t_n)(s - t_n)} w(\mathbf{x}, t_n) d\mathbf{x} \right|$$

$$\leq 2 \int_s^{t_n} \int_{\mathbb{R}^d} |f_\tau(\mathbf{z}, r)| \, |w(\mathbf{x}, t_n)| d\mathbf{x} dr$$

$$\leq 2\sqrt{2} (t_n - s)^{1/2} \|f_\tau\|_{L^2([s, t_n]; L^2(\mathbb{R}^d))} \|w(\mathbf{x}, t_n)\|_{L^2(\mathbb{R}^d)}.$$

Therefore,

$$(4.20) \qquad |I_f^{(3)}| \leq \frac{2\sqrt{2}}{3} \left[(\Delta t)^2 \|f_\tau\|_{L^2(J_n; L^2(\mathbb{R}^d))}^2 + \Delta t \|w(\mathbf{x}, t_n)\|_{L^2(\mathbb{R}^d)}^2 \right].$$

Now we piece together the above estimates (4.18), (4.19), and (4.20) to obtain an estimate on the source term

$$(4.21) \qquad \begin{aligned} |E(f, w)| &\leq C\Delta t \|w(\mathbf{x}, t_n)\|_{L^2(\mathbb{R}^d)}^2 \\ &\quad + C(\Delta t)^2 \left[\|f\|_{L^2(J_n; L^2(\mathbb{R}^d))}^2 + \|f_\tau\|_{L^2(J_n; L^2(\mathbb{R}^d))}^2 \right], \end{aligned}$$

where the constant C can be taken as $C = \frac{2\sqrt{2}}{\sqrt{3}} [\|\mathrm{div}\mathbf{V}\|_{\infty} + \|R\|_{\infty} + 1]$.

4.3. Final estimate. Combining estimates (4.15) and (4.21) with (4.4) and taking $w(\mathbf{x}, t_n)$ as $P_h u_n - U_h^n$, we get

$$\|P_h u_n - U_h^n\|_{L^2(\mathbb{R}^d)}^2 \leq \left(\frac{1}{2} + C\Delta t \right) \left(\|P_h u_n - U_h^n\|_{L^2(\mathbb{R}^d)}^2 + \|u_{n-1} - U_h^{n-1}\|_{L^2(\mathbb{R}^d)}^2 \right)$$

$$+ C(\Delta t)^3 \|u_{n-1}\|_{L^2(\mathbb{R}^d)}^2 + C(\Delta t)^2 \left[\|f\|_{L^2(J_n; L^2(\mathbb{R}^d))}^2 + \|f_\tau\|_{L^2(J_n; L^2(\mathbb{R}^d))}^2 \right].$$

Note that $\|u_{n-1} - U_h^{n-1}\|^2 = \|u_{n-1} - P_h u_{n-1}\|^2 + \|P_h u_{n-1} - U_h^{n-1}\|^2$. Taking Δt small enough such that $C\Delta t \leq 1/2$, we obtain

$$\|P_h u_n - U_h^n\|^2 \leq \left(\frac{1}{2} + C\Delta t\right)\left(\|P_h u_n - U_h^n\|^2 + \|P_h u_{n-1} - U_h^{n-1}\|^2\right)$$
$$+\|u_{n-1} - P_h u_{n-1}\|^2 + C(\Delta t)^3 \|u_{n-1}\|^2$$
$$+C(\Delta t)^2 \left[\|f\|^2_{L^2(J_n; L^2(\mathbb{R}^d))} + \|f_\tau\|^2_{L^2(J_n; L^2(\mathbb{R}^d))}\right].$$

Summing both sides for $n = 1, 2, \ldots, m$ ($m \leq N$) and canceling identical terms (also $U_h^0 = P_h u_0$), we get

$$\|P_h u_m - U_h^m\|^2 \leq \left(\frac{1}{2} + C\Delta t\right)\|P_h u_m - U_h^m\|^2 + C\Delta t \sum_{n=1}^{m-1} \|P_h u_n - U_h^n\|^2$$
$$+ \sum_{n=0}^{m-1} \|u_n - P_h u_n\|^2 + C(\Delta t)^3 \sum_{n=0}^{m-1} \|u_n\|^2$$
$$+C(\Delta t)^2 \left[\|f\|^2_{L^2(0,T; L^2(\mathbb{R}^d))} + \|f_\tau\|^2_{L^2(0,T; L^2(\mathbb{R}^d))}\right].$$

Taking Δt small enough such that $C\Delta t \leq 1/4$, we have

$$\|P_h u_m - U_h^m\|^2 \leq C\Delta t \sum_{n=1}^{m-1} \|P_h u_n - U_h^n\|^2$$
$$+C_0 \sum_{n=0}^{m-1} \|u_n - P_h u_n\|^2 + C(\Delta t)^3 \sum_{n=0}^{m-1} \|u_n\|^2$$
$$+C(\Delta t)^2 \left[\|f\|^2_{L^2(0,T; L^2(\mathbb{R}^d))} + \|f_\tau\|^2_{L^2(0,T; L^2(\mathbb{R}^d))}\right].$$

Applying the discrete Gronwall inequality, we obtain

$$(4.22) \quad \|P_h u_m - U_h^m\|^2_{L^2(\mathbb{R}^d)} \leq C \sum_{n=0}^{m-1} \|u_n - P_h u_n\|^2_{L^2(\mathbb{R}^d)} + C(\Delta t)^3 \sum_{n=0}^{m-1} \|u_n\|^2_{L^2(\mathbb{R}^d)}$$
$$+C(\Delta t)^2 \left[\|f\|^2_{L^2(0,T; L^2(\mathbb{R}^d))} + \|f_\tau\|^2_{L^2(0,T; L^2(\mathbb{R}^d))}\right].$$

The first two terms on the right side of the above estimate reflect the error buildup during the iterative process in numerical scheme (2.10). However, this can be controlled through the stability of the exact solution.

For the IVP to the linear convection-reaction equation in conservative form (1.1), Lemmas 3 and 4 in section 5 hold under the conditions in Assumption B and yield

$$(4.23) \quad \max_{0 \leq n \leq N} \|u_n\|_{L^2(\mathbb{R}^d)} \leq C\left(\|u_0\|_{L^2(\mathbb{R}^d)} + \|f\|_{L^2(0,T; L^2(\mathbb{R}^d))}\right),$$

$$(4.24) \quad \max_{0 \leq n \leq N} \|u_n\|_{H^1(\mathbb{R}^d)} \leq C\left(\|u_0\|_{H^1(\mathbb{R}^d)} + \|f\|_{L^2(0,T; H^1(\mathbb{R}^d))}\right),$$

where C depends only on the final time T and the norms of $\mathbf{V}, \mathrm{div}\mathbf{V}, R$ in the corresponding spaces in Assumption B.

Applying (4.23), we get

$$(4.25) \quad \sum_{n=0}^{m-1} \|u_n\|^2_{L^2(\mathbb{R}^d)} \leq \frac{C}{\Delta t}\left(\|u_0\|^2_{L^2(\mathbb{R}^d)} + \|f\|^2_{L^2(0,T; L^2(\mathbb{R}^d))}\right).$$

Combining Assumption A with (4.24), we obtain

$$(4.26) \qquad \sum_{n=0}^{m-1} \|u_n - P_h u_n\|_{L^2(\mathbb{R}^d)}^2 \leq C \frac{h^2}{\Delta t} \big(\|u_0\|_{H^1(\mathbb{R}^d)}^2 + \|f\|_{L^2(0,T;H^1(\mathbb{R}^d))}^2 \big).$$

Combining (3.1), (4.3), (4.22), and the above two estimates, we finish the proof of Theorem 1.   □

5. Stability of exact solutions. In this section, we prove the stability lemmas used in the last section about the exact solution to a linear convection-reaction equation. All results are first established for the solution to an IVP to the linear transport equation in nonconservative form:

$$(5.1) \qquad \begin{cases} u_t + \mathbf{V} \cdot \nabla u = cu + f, & (\mathbf{x}, t) \in \mathbb{R}^d \times (0, T], \\ u(\mathbf{x}, 0) = u_0(\mathbf{x}), & \mathbf{x} \in \mathbb{R}^d, \end{cases}$$

where $u_0 \in L^2(\mathbb{R}^d)$. However, the results can be easily passed to the linear convection-reaction equation in the conservative form (1.1) by simply setting $c = -(\mathrm{div}\mathbf{V} + R)$.

In this section, Assumption B, part (i), can be relaxed to $\mathbf{V} \in L^1(0, T; (W_\infty^1(\mathbb{R}^d))^d)$; that is, $\mathbf{V}$ is Lipschitz with respect to only spatial variables. Although it is possible to prove the existence, uniqueness, and regularity results below under even weaker assumptions, we restrict our attention to the case $\mathbf{V} \in L^1(0, T; (W_\infty^1(\mathbb{R}^d))^d)$ to keep the assumptions simple.

LEMMA 2 (existence and uniqueness). *Suppose* $\mathbf{V} \in L^1(0, T; (W_\infty^1(\mathbb{R}^d))^d)$, $c \in L^1(0, T; L^\infty(\mathbb{R}^d))$, *and* $f \in L^1(0, T; L^2(\mathbb{R}^d))$. *If* $u_0 \in L^2(\mathbb{R}^d)$, *then there exists a unique solution to* (5.1) *in* $L^\infty(0, T; L^2(\mathbb{R}^d))$ *and the solution can be explicitly expressed as*

$$(5.2) \quad u(\mathbf{x}, t) = u_0(\mathbf{y}(0; \mathbf{x}, t)) e^{\int_0^t c(\mathbf{y}(s; \mathbf{x}, t), s) ds} + \int_0^t f(\mathbf{y}(s; \mathbf{x}, t), s) e^{\int_s^t c(\mathbf{y}(r; \mathbf{x}, t), r) dr} ds.$$

The uniqueness of the solution u in Lemma 2 is a corollary of the results in [16]. The solution formula (5.2) can be derived using the results in [16], the techniques in [22], and standard density arguments.

LEMMA 3 (L^2-stability). *Assume that* $\mathbf{V}, c,$ *and* f *satisfy the conditions of Lemma 2. If* $u_0 \in L^2(\mathbb{R}^d)$, *then the unique solution satisfies*

$$(5.3) \qquad \|u(\cdot, t)\|_{L^2(\mathbb{R}^d)} \leq C \left[\|u_0\|_{L^2(\mathbb{R}^d)} + \|f\|_{L^1(0,t;L^2(\mathbb{R}^d))} \right] \qquad \forall t \in [0, T],$$

where C *can be taken as* $C = e^{\int_0^t [\|c(\cdot, r)\|_\infty + \frac{1}{2} \|\mathrm{div}\mathbf{V}(\cdot, r)\|_\infty] dr}$.

Proof. Applying the triangle and Minkowski inequalities to the representation formula (5.2), we obtain

$$(5.4) \qquad \begin{aligned} \|u(\cdot, t)\|_2 \leq e^{\int_0^t \|c(\cdot, s)\|_\infty ds} \Bigg[&\left(\int_{\mathbb{R}^d} |u_0(\mathbf{y}(0; \mathbf{x}, t))|^2 d\mathbf{x} \right)^{\frac{1}{2}} \\ &+ \int_0^t \left(\int_{\mathbb{R}^d} |f(\mathbf{y}(s; \mathbf{x}, t), s)|^2 d\mathbf{x} \right)^{\frac{1}{2}} ds \Bigg]. \end{aligned}$$

Let $\mathbf{J}(\mathbf{y}, \mathbf{x})$ be the Jacobian of mapping $\mathbf{x} \mapsto \mathbf{y} := \mathbf{y}(s; \mathbf{x}, t)$. It is known [13] that

$$(5.5) \qquad \mathbf{J}(\mathbf{y}, \mathbf{x}) = e^{\int_t^s \mathrm{div}\mathbf{V}(\mathbf{y}(r; \mathbf{x}, t), r) dr}.$$

Hence, for any $s \leq t \in [0, T]$, we have

$$(5.6) \qquad e^{-\int_s^t \|\mathrm{div}\mathbf{V}(\cdot, r)\|_\infty dr} \leq \mathbf{J}(\mathbf{y}, \mathbf{x}) \leq e^{\int_s^t \|\mathrm{div}\mathbf{V}(\cdot, r)\|_\infty dr}.$$

Of course, we can take $\mathbf{y} := \mathbf{y}(s; \mathbf{x}, t)$ as the starting point. Then the reversibility of flow ensures that the Jacobian $\mathbf{J}(\mathbf{x}, \mathbf{y})$ of the inverse mapping $\mathbf{y}(s; \mathbf{x}, t) \mapsto \mathbf{x}$ satisfies the same estimate. A change of variables in (5.4) and then an application of the above estimates on Jacobians conclude the proof. $\square$

Next we shall impose some additional conditions on c and f so that the solution $u \in L^\infty(0, T; H^1(\mathbb{R}^d))$, provided $u_0 \in H^1(\mathbb{R}^d)$. In other words, under these sufficient conditions on $\mathbf{V}$, c, and f, the solution operator $E_t : u_0 \mapsto u(\cdot, t)$ is a bounded operator from $H^1(\mathbb{R}^d)$ to $H^1(\mathbb{R}^d)$.

LEMMA 4 (H^1-stability). *Assume that* $\mathbf{V} \in L^1(0, T; (W_\infty^1(\mathbb{R}^d))^d)$, $c \in L^1(0, T; W_\infty^1(\mathbb{R}^d))$, *and* $f \in L^1(0, T; H^1(\mathbb{R}^d))$. *If* $u_0 \in H^1(\mathbb{R}^d)$, *then the unique weak solution satisfies*

$$(5.7) \qquad \|u(\cdot, t)\|_{H^1(\mathbb{R}^d)} \leq C\big(\|u_0\|_{H^1(\mathbb{R}^d)} + \|f\|_{L^1(0, t; H^1(\mathbb{R}^d))}\big) \quad \forall t \in [0, T],$$

where the constant C *depends only on* $\|\mathbf{V}\|_{L^1(0, T; (W_\infty^1(\mathbb{R}^d))^d)}$ *and* $\|c\|_{L^1(0, T; W_\infty^1(\mathbb{R}^d))}$.

Proof. Using the chain rule to the solution formula (5.2), we can derive expressions for all partial derivatives $\frac{\partial u}{\partial x_i}, 1 \leq i \leq d$, which involve spatial partial derivatives of the flow $\mathbf{y}(s; \mathbf{x}, t), c$, and f. Let $\mathbf{z}(s; \mathbf{x}, t)$ be any partial derivative of $\mathbf{y} := \mathbf{y}(s; \mathbf{x}, t)$ with respect to one of the variables $x_1, \ldots, x_d$. We shall derive a uniform bound for $\mathbf{z}(s; \cdot, t)$. Let $|\mathbf{z}(s; \mathbf{x}, t)|$ be the usual Euclidean norm for a vector in $\mathbb{R}^d$. The following assertion was proven in [8, 22]: Let $\mathbf{V}$ satisfy

$$|(\mathbf{V}(\mathbf{x}, t) - \mathbf{V}(\mathbf{y}, t)) \cdot (\mathbf{x} - \mathbf{y})| \leq K(t)|\mathbf{x} - \mathbf{y}|^2;$$

then

$$|\mathbf{z}(s; \mathbf{x}, t)| \leq |\mathbf{z}(t; \mathbf{x}, t)| e^{\int_0^T K(r) dr} = e^{\int_0^T K(r) dr}.$$

Here we have used the fact $\mathbf{y}(t; \mathbf{x}, t) = \mathbf{x}$ to get $|\mathbf{z}(t; \mathbf{x}, t)| = 1$. It is not difficult to verify that if $\mathbf{V} \in L^1(0, T; (W_\infty^1(\mathbb{R}^d))^d)$, then the above inequality holds with $K(t) = \|\mathbf{V}(\cdot, t)\|_{W_\infty^1(\mathbb{R}^d)}$. Therefore,

$$(5.8) \qquad \|\mathbf{z}(s, \cdot, t)\|_\infty := \mathrm{ess\,sup}_{\mathbf{x} \in \mathbb{R}^d} |\mathbf{z}(s, \mathbf{x}, t)| \leq e^{\|\mathbf{V}\|_{L^1(0, T; (W_\infty^1(\mathbb{R}^d))^d)}}.$$

This is the standard estimate for a Lipschitz flow; see (1.4) in [7]. Based on the uniform bound (5.8), we can estimate all partial derivatives as follows:

$$\left\|\frac{\partial u}{\partial x_i}(\cdot, t)\right\|_{L^2(\mathbb{R}^d)} \leq C\big(\|u_0\|_{H^1(\mathbb{R}^d)} + \|f\|_{L^1(0, T; H^1(\mathbb{R}^d))}\big),$$

where the constant C depends only on the norms of $\mathbf{V}$ in $L^1(0, T; (W_\infty^1(\mathbb{R}^d))^d)$ and c in $L^1(0, T; W_\infty^1(\mathbb{R}^d))$. We finish the proof by combining the above estimate and the result of Lemma 3. $\square$

6. Extension to Besov spaces. Besov spaces provide subtle characterization of regularity of functions. Interpolation of spaces and operators is a classical topic in harmonic analysis and has many interesting applications. In this section, we cite only

the minimal requisite for our discussion. Readers are referred to [3, 15] for complete accounts of these two topics.

Besov spaces involve moduli of smoothness rather than the distributional derivatives used for Sobolev spaces. Let $1 \leq p \leq \infty$, $f \in L^p(\mathbb{R}^d)$, $\mathbf{h} \in \mathbb{R}^d$, and let $\Delta_{\mathbf{h}}$ be the usual difference operator: $\Delta_{\mathbf{h}} f(\mathbf{x}) := f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})$. Let k be a positive integer and $t > 0$. The kth modulus of smoothness of function f is defined as

$$(6.1) \qquad \omega_k(f, t)_p := \sup_{|\mathbf{h}| \leq t} \|\Delta_{\mathbf{h}}^k(f, \mathbf{x})\|_{L^p(\mathbb{R}^d)}.$$

Suppose $\alpha > 0$ and $0 < q \leq \infty$. Let k be a positive integer such that $k > \alpha$. The Besov space $B_{p,q}^{\alpha}(\mathbb{R}^d)$ consists of functions $f \in L^p(\mathbb{R}^d)$ (if $p < \infty$) or $C(\mathbb{R}^d)$ (if $p = \infty$) such that

$$(6.2) \qquad \|f\|_{B_{p,q}^{\alpha}} := \begin{cases} \|f\|_{L^p} + \left\{ \int_0^{\infty} \left[t^{-\alpha} \omega_k(f, t)_p \right]^q \dfrac{dt}{t} \right\}^{1/q}, & 0 < q < \infty, \\ \|f\|_{L^p} + \sup_{t > 0} t^{-\alpha} \omega_k(f, t)_p, & q = \infty, \end{cases}$$

is finite.

It is known that $B_{p,q_1}^{\alpha} \subset B_{p,q_2}^{\alpha}$ if $q_1 < q_2$. When $p = 2$, one has $B_{2,2}^{\alpha} = H^{\alpha}$ with equivalent norms.

Interpolation of spaces can be defined by K-functionals. Let $X_1 \subset X_0$ be Banach spaces. For any $f \in X_0$ and $t > 0$,

$$(6.3) \qquad K(f, t) := K(f, t; X_0, X_1) := \inf_{g \in X_1} \{\|f - g\|_{X_0} + t\|g\|_{X_1}\}.$$

Here t is viewed as a penalty factor. The intermediate space $[X_0, X_1]_{\theta, q}$ consists of all $f \in X_0$ for which

$$(6.4) \qquad \|f\|_{\theta, q} := \begin{cases} \left\{ \int_0^{\infty} \left[t^{-\theta} K(f, t) \right]^q \dfrac{dt}{t} \right\}^{1/q}, & 0 < q < \infty, \\ \sup_{t > 0} t^{-\theta} K(f, t), & q = \infty, \end{cases}$$

is finite. Obviously, $X_1 \subset [X_0, X_1]_{\theta, q} \subset X_0$.

Amazingly, the K-functional for the pair (L^p, W_p^k) is equivalent to the modulus of smoothness, and hence we have the following (see [3, 15]).

LEMMA 5 (interpolation of spaces). *Let $k > 0$ be an integer and let $1 \leq p \leq \infty$. For any $0 < \theta < 1, 0 < q \leq \infty$, we have*

$$(6.5) \qquad \left[L^p(\mathbb{R}^d), W_p^k(\mathbb{R}^d) \right]_{\theta, q} = B_{p,q}^{\theta k}(\mathbb{R}^d).$$

Especially,

$$(6.6) \qquad \left[L^2(\mathbb{R}^d), H^1(\mathbb{R}^d) \right]_{\theta, q} = B_{2,q}^{\theta}(\mathbb{R}^d).$$

LEMMA 6 (interpolation of operators). *Suppose that $X_1 \subset X_0$ and Y are Banach spaces. If T is a linear operator from X_i to Y with norm M_i $(i = 0, 1)$, then T is also a linear operator from $[X_0, X_1]_{\theta, q}$ to Y with norm not exceeding $M_0^{1-\theta} M_1^{\theta}$ for any $0 < \theta < 1, 0 < q \leq \infty$.*

THEOREM 7. *Suppose that Assumptions* A *and* B *are satisfied. Then for any* $0 < \theta < 1, 0 < q \le \infty$, *the following error estimate holds:*

$$
\max_{0 \le n \le N} \|u(\mathbf{x}, t_n) - U_h^n(\mathbf{x})\|_{L^2(\mathbb{R}^d)}
$$

$$
(6.7) \qquad \le C\Big\{ \Big[(h/\sqrt{\Delta t})^\theta + (\Delta t)^\theta\Big] \|u_0\|_{B_{2,q}^\theta(\mathbb{R}^d)}
$$

$$
+ (h/\sqrt{\Delta t} + \Delta t) \Big[\|f\|_{L^2(0,T;H^1(\mathbb{R}^d))} + \|f_\tau\|_{L^2(0,T;L^2(\mathbb{R}^d))}\Big]\Big\},
$$

where Δt *and* h *bear the same meaning as that in Theorem* 1, *but* C *may additionally depend on* θ.

Proof. We split the error as

$$
u_n - U_h^n = [u_n^{(1)} - U_h^{(1),n}] + [u_n^{(2)} - U_h^{(2),n}],
$$

where $u_n^{(1)}$ and $U_h^{(1),n}$ are the exact and numerical solutions for problem (1.1) without source term (i.e., $f \equiv 0$), whereas $u_n^{(2)}$ and $U_h^{(2),n}$ are the exact and numerical solutions for the problem with no initial data ($u_0 \equiv 0$).

Recall that in [24] we proved the numerical solution is L^2-stable. Here (4.23) is the L^2-stability of the exact solution. Combined, they imply

$$
\|u_n^{(1)} - U_h^{(1),n}\|_{L^2(\mathbb{R}^d)} \le C \|u_0\|_{L^2(\mathbb{R}^d)}
$$

for $u_0 \in L^2(\mathbb{R}^d)$. If $u_0 \in H^1(\mathbb{R}^d)$, then by Theorem 1 we have

$$
\|u_n^{(1)} - U_h^{(1),n}\|_{L^2(\mathbb{R}^d)} \le C(h/\sqrt{\Delta t} + \Delta t)\|u_0\|_{H^1(\mathbb{R}^d)}.
$$

Applying Lemmas 5 and 6 to the linear operator $E_n : u_0 \mapsto u_n^{(1)} - U_h^{(1),n}$, we obtain

$$
\|u_n^{(1)} - U_h^{(1),n}\|_{L^2(\mathbb{R}^d)} \le C(h/\sqrt{\Delta t} + \Delta t)^\theta \|u_0\|_{B_{2,q}^\theta(\mathbb{R}^d)}
$$

$$
\le C\Big[(h/\sqrt{\Delta t})^\theta + (\Delta t)^\theta\Big]\|u_0\|_{B_{2,q}^\theta(\mathbb{R}^d)}.
$$

By Theorem 1 again, we have

$$
\|u_n^{(2)} - U_h^{(2),n}\|_{L^2(\mathbb{R}^d)} \le C(h/\sqrt{\Delta t} + \Delta t)\big(\|f\|_{L^2(0,T;H^1(\mathbb{R}^d))} + \|f_\tau\|_{L^2(0,T;L^2(\mathbb{R}^d))}\big).
$$

Then the conclusion of Theorem 7 follows from a triangle inequality. ☐

7. Optimality of our error estimates. For solutions with H^1- or even lower regularity, we can use Haar wavelets to carry out our numerical approximations. Haar wavelets are the simplest wavelets and have only one vanishing moment. We know that the order of approximation accuracy is usually the minimum of the order of smoothness of the function being approximated and the order of the method, which is the number of vanishing moments for orthogonal wavelets. For a solution $u \in H^1$, Assumption A in section 3 is satisfied for the approximation subspace V_h generated by Haar wavelets.

However, in the original formulation of ELLAM discussed in subsection 2.1, the test function $w(\mathbf{x}, t)$ is assumed to be in $H^1(\mathbb{R}^d)$ for any $t \in (t_{n-1}, t_n]$ and required to satisfy the adjoint equation (2.2). But Haar scaling functions and wavelets are not in $H^1(\mathbb{R})$. However, mollifications with any cut-off function can be applied to Haar

scaling functions or wavelets [18], so that (2.2) is satisfied for the mollifications. Then, based on the L^2-stability of the exact solution and density arguments, we can still get the reference equation (2.5) when the test function $w(x,t)$ is such that $w(x,t_n)$ is taken as a Haar scaling function or wavelet. Numerical scheme (2.10) can be established for Haar scaling functions and wavelets without difficulty. Moreover, when Haar scaling functions are used in the ELLAM formulation, the coefficients are exactly the cell averages of the unknown function on dyadic cells, thus we have local conservation of mass.

Let us consider a simple convection equation: $u_t - u_x = 0$, with a time step $\Delta t = \alpha h$ ($0 < \alpha < 1$). For this case, the ELLAM scheme with Haar scaling functions becomes a monotone scheme, which is widely used for hyperbolic conservation laws. Monotone schemes are a special case of linear formal first order schemes; see [4] for details. According to Theorem 4.4 in [4], there exists a constant C_1 such that

$$(7.1) \qquad \sup_{\|u_0\|_{H^1(\mathbb{R})} \leq 1} \|u(\mathbf{x}, T) - U_h^N(\mathbf{x})\|_{L^2(\mathbb{R})} \geq C_1 h^{1/2}.$$

An explicit value for C_1 can be derived using a modification of the argument in [23].

On the other hand, for this equation with any initial condition $u_0 \in H^1(\mathbb{R})$ and $\Delta t = \alpha h$ ($0 < \alpha < 1$), our Theorem 1 implies an upper bound for the error in the numerical solution at the final time T:

$$(7.2) \qquad \|u(\mathbf{x}, T) - U_h^N(\mathbf{x})\|_{L^2(\mathbb{R})} \leq C_2 h^{1/2}.$$

The above lower and upper bounds imply that our error estimate is optimal for this case for the class of initial conditions $u_0 \in H^1$. We also want to point out that it is possible to give more examples with an optimal rate $1/2$: for a different equation or a different subspace generated by smooth basis functions.

Regarding the term $h/\sqrt{\Delta t}$ in the error estimates, our understanding is that it reflects the behavior of numerical schemes when only H^1-smoothness is assumed for exact solutions. From approximation theory, we know that only first order $\mathcal{O}(h)$ approximation accuracy can be achieved at each fixed time step. For time-marching schemes, error will accumulate. However, the buildup of error can be controlled by the stability of solutions. In our theoretical estimates, we derive an upper bound for the error in the form $h/\sqrt{\Delta t}$. A similar error estimate with adverse dependence on Δt in the form

$$\|u - u_h\|_{L^\infty(0,T;L^1(\mathbb{R}))} \leq \|u_0\|_{TV} \left(h + \frac{2\sqrt{T}}{\sqrt{3}} \frac{h}{\sqrt{\Delta t}} \right)$$

can also be found in [20] for one-dimensional nonlinear conservation law. Here u_h is the numerical solution and $\| \cdot \|_{TV}$ denotes the total variation.

To balance the two parts in our error estimates, the optimal choice for Δt is $\Delta t = Ch^{2/3}$. No lower bound for the error for the case $\Delta t = Ch^\beta$ ($\beta \neq 1$) is covered in [4, 23]. It is also almost impossible to derive a general error estimate that is optimal for all schemes and all $0 < \beta < \infty$. But the rates observed in our numerical experiments in the next section are close to our theoretical rates for the case $\Delta t = Ch^{2/3}$.

It might not be exciting if small time steps Δt have to be used, e.g., for the traditional finite difference methods that are subject to the CFL condition. However, our ELLAM schemes are CFL-free [24]. So we are allowed to use relatively large time steps. This saves computations while retaining accuracy since information from characteristics are exploited.

8. Numerical experiments. In this section, we present one-dimensional numerical experiments to illustrate the theoretical results proven in previous sections.

We consider two examples with $V(x,t) = 1$, $R(x,t) = 0.2\sin t$, $f(x,t) = 0$, $\Omega = [0,2]$, and $T = 1$. The exact solution is then given by $u(x,t) = u_0(\xi)e^{0.2(\cos t - 1)}$, where ξ is obtained by back-tracking the characteristic from (x,t) to $(\xi,0)$.

Example 1. The initial condition is specified as a cusp function:

$$(8.1) \qquad u_0(x) = \begin{cases} A\left(1 - \left|\dfrac{x-\alpha}{\beta}\right|^{\frac{1}{2}+\gamma}\right) & \text{if } |x-\alpha| \le \beta, \\ 0 & \text{otherwise,} \end{cases}$$

where $\alpha \in \mathbb{R}, \beta > 0$, and $0 < \gamma \le 1/2$. It can be verified that $u_0 \in H^{1+\delta}(\mathbb{R})$ for any $0 < \delta < \gamma$.

In the numerical experiment, we take $A = 1, \alpha = 0.5, \beta = 0.25, \gamma = 0.01$. So the initial data is barely in $H^1(\mathbb{R})$, or, precisely, $u_0 \in H^{1+\delta}$ for $0 < \delta < 0.01$. The second order Daubechies scaling function and wavelet are used in the wavelet schemes.

We can attain only first order approximation in space for the initial data since it is barely in $H^1(\mathbb{R})$. The adverse dependence of the error estimate on Δt in section 3 indicates that time discretization has to be done carefully. More time steps do not necessarily mean better approximations. The best result can be attained when $\Delta t = h/\sqrt{\Delta t}$, that is, $\Delta t = h^{2/3}$, if the constants in the estimate are ignored. In other words, the wavelet schemes have convergence rate $h^{2/3}$. In Table 8.1, the numerical solution at the final time step has a convergence rate 0.74 in h, which is just about 10% better than the theoretical estimate $2/3$.

TABLE 8.1
Convergence rates in h for Example 1.

h	$\Delta t = Ch^{2/3}$	$\|u_0 - U_0\|_{L^2}$	$\|u_T - U_T\|_{L^2}$
$1/2^6$	$1/16$	7.502E-3	5.134E-3
$1/2^9$	$1/64$	9.181E-4	1.114E-4
$1/2^{12}$	$1/256$	1.124E-4	2.457E-4
$1/2^{15}$	$1/1024$	1.376E-5	4.880E-5
Convergence rates		1.01	0.74

Example 2. The initial condition is the indicator function $\chi_{[\alpha,\beta]}$ of the interval $[\alpha,\beta]$:

$$(8.2) \qquad u_0(x) = \chi_{[\alpha,\beta]} = \begin{cases} 1 & x \in [\alpha,\beta], \\ 0 & \text{otherwise.} \end{cases}$$

It is known that, for any interval I containing $[\alpha,\beta]$ in its interior, we have $\chi_{[\alpha,\beta]} \in H^{\frac{1}{2}-\delta}(I)$ for any $0 < \delta \le \frac{1}{2}$ but $\notin H^{\frac{1}{2}}(I)$. Direct calculations indicate that

$$(8.3) \qquad \|\chi_{[\alpha,\beta]}\|_{H^{\frac{1}{2}-\delta}(I)} = \mathcal{O}(\delta^{-\frac{1}{2}}) \to \infty \quad \text{as } \delta \to 0.$$

Hence, the approximation order could not be characterized well using norms in Sobolev spaces. This is one place where we should use Besov spaces. It can be verified that $\chi_{[\alpha,\beta]} \in B_{2,\infty}^{\frac{1}{2}}(I)$ but $\notin B_{2,q}^{\frac{1}{2}}(I)$ for $q < \infty$.

The numerical results for $[\alpha,\beta] = [0.25, 0.75]$ are shown in Table 8.2. The order of approximation to the initial data (u_0) is exactly $1/2$ because $u_0 \in B_{2,\infty}^{\frac{1}{2}}$. After the

TABLE 8.2
Convergence rates in h for Example 2.

h	$\Delta t = Ch^{2/3}$	$\|u_0 - U_0\|_{L^2}$	$\|u_T - U_T\|_{L^2}$
$1/2^7$	$1/20$	3.232E-2	5.375E-2
$1/2^{10}$	$1/80$	1.142E-2	2.248E-2
$1/2^{13}$	$1/320$	4.040E-3	1.151E-2
$1/2^{16}$	$1/1280$	1.428E-3	5.053E-3
Convergence rates		0.50	0.38

time-marching procedure, the approximation error at the final time step is of order 0.38, close to the theoretical estimate 1/3.

In these two numerical examples, convergence rates are a little better than the theoretically proven rates because the velocity field is nice (but we can easily compute the exact solutions in these cases). If the velocity field is exactly Lipschitz, i.e., it has minimal regularity, we expect numerical results to be even closer to the theoretical estimates.

9. Concluding remarks. Some similar results concerning convergence rates of Lagrangian–Galerkin methods for convection-dominated diffusion problems are presented in [2]. Their estimates are uniform in the small diffusion parameter and can be carried over to the hyperbolic limit case—linear convection equations without a reaction term. Their results are consistent with ours but require a smoother velocity field ($\mathbf{V} \in C(0, T; C^2(\overline{\Omega}))$).

The assumption $R \in W^1_\infty(\mathbb{R}^d \times [0, T])$ in our paper means some smoothness in the reaction term is required. Note that the test functions in ELLAM rely on the properties of the reaction along characteristics. Generally speaking, this type of smoothness is needed for time-marching schemes; otherwise we could not use well the information of the solutions at previous time steps.

In section 6, we did not discuss interpolations on the source term. The limitation is due to the way the source term is truncated in numerical scheme (2.10). Recall that in (2.7) we approximate the double integral for the source term by a single integral. This assumes some smoothness of the source term along the temporal direction or the characteristic direction. In return, the computational cost for the source term is reduced. Of course, the numerical scheme can be modified to allow even lower regularity for the source term, but accordingly the computational cost will increase.

For the wavelet scheme with adaptive compression, i.e., Scheme III in [24], computational cost will be significantly reduced through thresholding wavelet coefficients in the smooth regions. But the approximation will deteriorate as the threshold parameter is increased. A good choice of the threshold, in other words, a quantitative description of the trade-off between computational cost and approximation accuracy, is a delicate issue of nonlinear approximation and is already under our investigation.

REFERENCES

[1] C. BACUTA, J. H. BRAMBLE, AND J. XU, *Regularity estimates for elliptic boundary value problems in Besov spaces*, Math. Comp., 72 (2003), pp. 1577–1595.

[2] M. BAUSE AND P. KNABNER, *Uniform error analysis for Lagrangian–Galerkin approximations of convection-dominated problems*, SIAM J. Numer. Anal., 39 (2002), pp. 1954–1984.

[3] C. BENNETT AND R. SHARPLEY, *Interpolation of Operators*, Academic Press, San Diego, CA, 1988.

[4] P. Brenner, V. Thomée, and, L. B. Wahlbin, *Besov Spaces and Applications to Difference Methods for Initial Value Problems*, Lecture Notes in Math. 434, Springer-Verlag, Berlin, 1975.

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

[6] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics Appl. Math. 40, SIAM, Philadelphia, 2002.

[7] F. Colombini and N. Lerner, *Uniqueness of continuous solutions for BV vector fields*, Duke Math. J., 111 (2002), pp. 357–384.

[8] E. Conway, *Generalized solutions of linear differential equations with discontinuous coefficients and the uniqueness question for multidimensional quasilinear conservation laws*, J. Math. Anal. Appl., 18 (1967), pp. 238–251.

[9] W. Dahmen, *Wavelet and multiscale methods for operator equations*, Acta Numer., 6 (1997), pp. 55–228.

[10] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 61, SIAM, Philadelphia, 1992.

[11] C. N. Dawson, T. F. Dupont, and M. F. Wheeler, *The rate of convergence of the modified method of characteristics for linear advection equation in one dimension*, in Mathematics for Large Scale Computing, J. C. Diaz, ed., Marcel Dekker, New York, 1989, pp. 115–126.

[12] 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.

[13] B. Desjardins, *A few remarks on ordinary differential equations*, Comm. Partial Differential Equations, 21 (1996), pp. 1667–1703.

[14] R. A. DeVore, *Nonlinear approximation*, Acta Numer., 7 (1998), pp. 51–150.

[15] R. A. DeVore and G. G. Lorentz, *Constructive Approximation*, Springer-Verlag, Berlin, 1993.

[16] R. DiPerna and P. Lions, *Ordinary differential equations, transport theory, and Sobolev spaces*, Invent. Math., 98 (1989), pp. 511–547.

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

[18] L. C. Evans, *Partial Differential Equations*, AMS, Providence, RI, 1998.

[19] R. E. Ewing and H. Wang, *An optimal-order estimate for Eulerian–Lagrangian localized adjoint methods for variable-coefficient advection-reaction problems*, SIAM J. Numer. Anal., 33 (1996), pp. 318–348.

[20] B. J. Lucier, *Error bounds for the methods of Glimm, Godunov, and LeVeque*, SIAM J. Numer. Anal., 22 (1985), pp. 1074–1081.

[21] K. W. Morton, A. Priestley, and E. Süli, *Stability of the Lagrange–Galerkin method with nonexact integration*, RAIRO Modél. Math. Anal. Numér., 22 (1988), pp. 625–653.

[22] G. Petrova and B. Popov, *Linear transport equations with μ-monotone coefficients*, J. Math. Anal. Appl., 260 (2001), pp. 307–324.

[23] T. Tang and Z. H. Teng, *The sharpness of Kuznetsov's $O(\sqrt{\Delta x})L^1$-error estimate for monotone difference schemes*, Math. Comput., 64 (1995), pp. 581–589.

[24] H. Wang and J. Liu, *Development of CFL-free, explicit schemes for multidimensional advection-reaction equations*, SIAM J. Sci. Comput., 23 (2001), pp. 1418–1438.

[25] H. Wang, X. Shi, and R. E. Ewing, *An ELLAM scheme for multidimensional advection-reaction equations and its optimal-order error estimate*, SIAM J. Numer. Anal., 38 (2001), pp. 1846–1885.