

A METHOD TO CALCULATE NUMERICAL ERRORS USING ADJOINT ERROR ESTIMATION FOR LINEAR ADVECTION

JEFFREY M. CONNORS^{†‡}, JEFFREY W. BANKS[†], JEFFREY A. HITTINGER[†], AND
CAROL S. WOODWARD[†]

Abstract. This paper is concerned with the computation of numerical discretization error for uncertainty quantification. An *a posteriori* error formula is described for a functional measurement of the solution to a scalar advection equation that is estimated by finite volume approximations. An exact error formula and computable error estimate are derived based on an abstractly defined approximation of the adjoint solution. The adjoint problem is divorced from the finite volume method used to approximate the forward solution variables and may be approximated using a low-order finite volume method. The accuracy of the computable error estimate provably satisfies an *a priori* error bound for sufficiently smooth solutions of the forward and adjoint problems. Computational examples are provided that show support of the theory for smooth solutions. The application to problems with discontinuities is also investigated computationally.

Key words. a posteriori error, adjoint error, error estimation, hyperbolic conservation law

1. Introduction. The problem is to calculate the error in an estimate of $\mathcal{M}(u)$, where $u : \mathbb{R}^D \times \mathbb{R}^+ \rightarrow \mathbb{R}$ solves

$$\partial_t u + \nabla \cdot (\vec{a}u) = F(x, t), \quad (x, t) \in \Omega_T, \quad (1.1)$$

$$u(x, 0) = u_0(x), \quad x \in \Omega \subset \mathbb{R}^D, \quad D \in \{1, 2, 3\}, \quad (1.2)$$

for a constant vector $\vec{a} = \langle a_1, \dots, a_D \rangle$ and where $\mathcal{M}(u) \in \mathbb{R}$ is a quantity of interest (QoI). Periodic boundary conditions are assumed in space. The spatial domain is a box,

$$\Omega = (x_1^L, x_1^R) \times \dots \times (x_D^L, x_D^R),$$

and the global space-time domain is denoted by $\Omega_T = \Omega \times (0, T]$. We assume the QoI may be expressed in one of these two forms:

$$\mathcal{M}(u) = \int_{\Omega_T} \psi(x, t) u(x, t) dx dt \quad (1.3)$$

or

$$\mathcal{M}(u) = \int_{\Omega} \psi(x) u(x, T) dx. \quad (1.4)$$

These functionals are linear with respect to u . The assumptions of linearity of the QoI and periodic boundary conditions are not essential, but serve to simplify the analysis in this report. The extension of the proposed techniques to accommodate nonlinear functionals of the solution and other boundary conditions depends on the functional, but in many cases will follow the approach outlined in [5, 6], assuming a sufficiently smooth solution of (1.1)-(1.2).

Assume that a finite volume method has been applied to estimate the solution of (1.1)-(1.2), yielding data denoted by \bar{v} . The QoI is estimated in two steps:

[†]Lawrence Livermore National Laboratory, Center for Applied Scientific Computing, Livermore, CA, 94551, USA. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 and was funded by the Uncertainty Quantification Strategic Initiative Laboratory Directed Research and Development Project at LLNL under project tracking code 10-SI-013. LLNL-JRNL-484291.

[‡]Email: connors4@llnl.gov.

1. Define a continuum reconstruction \tilde{u} from the finite volume data set \bar{v} .
2. Calculate $\mathcal{M}(\tilde{u})$; this is the approximation of $\mathcal{M}(u)$.

In large scale simulations it is expected that the discretization error in $\mathcal{M}(\tilde{u})$ is significant. There has been increasing interest in performing uncertainty quantification studies to help understand the meaning of a QoI from a simulation, which is necessary to make informed decisions based on the code output. Calculating discretization errors in the QoI is necessary for numerous applications where these errors will comprise a significant component of the uncertainty in the QoI.

In practice, no method exists to calculate the error $\mathcal{M}(u) - \mathcal{M}(\tilde{u})$ exactly, since the true solution is not known. There are a number of techniques for estimating the error, (see, e.g., [28, 31] for a review), one of which is adjoint error estimation. This technique requires deriving an auxiliary (adjoint) problem for the error in the QoI. Then the error in the QoI can be calculated with a precision limited by the numerical error in approximating the adjoint solution. A theoretical justification has been provided for a variety of numerical methods for elliptic and parabolic problems, (see, e.g., [7, 13–16, 21, 27] and references therein).

Numerous numerical studies exist for application to hyperbolic conservation laws, such as [3–6, 8, 17, 19, 24–26, 32]. Many of these techniques exhibit good error estimation properties, but often, no rigorous theoretical justifications are provided due to the challenges posed by hyperbolic conservation laws, e.g. accommodating weak solutions with discontinuities or performing a linearization step in deriving the adjoint problem for the error. The reader is referred to [27] for some introduction to adjoint theory within the context of hyperbolic conservation laws. A review of adjoint-based error estimation techniques can be found in the review papers of Giles and Süli [21] as well as Fidkowski and Darmofal [18].

Many codes are sophisticated enough that modifying them or even understanding entire calculations can translate into an immense amount of work. Thus an emphasis is placed on developing a framework for *a posteriori* error estimation that limits intrusiveness. This report describes such a technique in Section 2, deriving an adjoint error estimate requiring little or no knowledge of the finite volume solver used to estimate the solution of the hyperbolic conservation law. Such an approach avoids the technical difficulties associated with linearizations, due to algorithmic nonlinearities, that have been present in some previous reports, [3–6, 24–26], and can clarify the issue of well-posedness of the adjoint problem. The definition of the adjoint problem depends on the forward problem, initial conditions, boundary conditions and QoI. Although previous reports have often been motivated by adaptive mesh refinement techniques to control discretization error, this report is instead motivated by the need for an efficient, robust, easily implemented error estimation technique to ascertain deterministic errors in computations for uncertainty quantification. Some discussion of efficiency is also provided, though this will not be the focus.

As a first step, the adjoint error estimation method used herein is introduced for the linear advection problem (1.1)-(1.2), in Section 2. An exact error representation formula and computable error estimate are derived in Section 2.3. The formulas require continuum approximations of the forward and adjoint solution variables, which are defined using a reconstruction process for finite volume data described in Section 3. This post-processing step is independent of details of the finite volume solvers such as nonlinearity of flux calculations. A similar approach was described in [29] using C^2 -cubic spline interpolation. In contrast, we use a localized and discretely conservative reconstruction procedure that is more consistent with standard finite volume

approaches and that applies in a natural way to the reconstruction of weak solutions. We derive rigorous error bounds and thereby justify an asymptotic statement of accuracy for the computable error estimate in Section 4. Based on this result, it is shown that if the true error in the estimated QoI exhibits an optimal or suboptimal rate of convergence, then the computable error estimate will converge to the true error faster than the true error converges to zero, using any first-order or better approximation of the adjoint solution. If the error in the estimated QoI converges at a superoptimal rate, a higher-order estimate of the adjoint solution may be required. In Section 5, the theory is explored in numerous computational examples, including a case involving a discontinuous adjoint solution. Conclusions and future work directions are discussed in Section 6.

2. The adjoint problem and error representation formula. The goal is to derive an exact formula for the error in a QoI. A continuous adjoint approach is taken (see e.g. Pierce and Giles [29]); the adjoint problem is determined by considering the adjoint PDE associated in the standard way with (1.1) and then determining appropriate forcing and boundary conditions given the desired QoI. It is assumed the forward and adjoint problems are approximated by some possibly discontinuous functions. In order to derive an error representation formula, one must account for the discontinuous nature of the approximations. Thus a bilinear form is identified with the forward problem over trial and test spaces containing the necessary class of approximating functions. A dual bilinear form is then derived and associated with the adjoint PDE problem in the strong form. Subsequently an exact error representation formula is derived. This approach is feasible for solutions with regularity $u \in H^1(\Omega_T)$, and this approach may easily be implemented with any finite volume method.

2.1. Preliminaries and notation. Denote a cellular partition of Ω with maximum cell width $h \in \mathbb{R}^+$ by \mathcal{K}^h . Let \mathcal{A} be a set of multi-indices $\alpha = (\alpha_1, \dots, \alpha_D) \in \mathbb{N}^D$ such that $\alpha_d \in \{1, 2, \dots, M_d\}$ for $d \in \{1, \dots, D\}$. The partition \mathcal{K}^h is the union of cells \mathcal{K}_α^h such that

$$\mathcal{K}_\alpha^h = \prod_{d=1}^D \left((x_d)_{\alpha_d} - \frac{1}{2}(h_d)_{\alpha_d}, (x_d)_{\alpha_d} + \frac{1}{2}(h_d)_{\alpha_d} \right), \quad (2.1)$$

where cell centers are denoted by x_α . For compactness, coordinate values at cell interfaces are denoted by $(x_d)_{\alpha_d \pm 1/2} = (x_d)_{\alpha_d} \pm \frac{1}{2}(h_d)_{\alpha_d}$ for $d = 1, \dots, D$. A function space with the necessary properties to later derive the *a posteriori* error formula is the broken space [6]

$$\mathcal{V}^B = \{v|_{\mathcal{K}_\alpha^h} \in H^1(\mathcal{K}_\alpha^h \times (0, T)), \forall \alpha \in \mathcal{A}\}. \quad (2.2)$$

Define a norm for all $v \in \mathcal{V}^B$ as the positive square root of

$$\|v\|_{\mathcal{V}^B}^2 = \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^h \times (0, T)} \left(|v(x, t)|^2 + |\partial_t v(x, t)|^2 + \sum_{d=1}^D |\partial_{x_d} v(x, t)|^2 \right) dx dt. \quad (2.3)$$

A trivial extension operator $T : v \in \mathcal{V}^B \rightarrow T(v) \in L^2(\Omega_T)$ exists, so for convenience the notation is suppressed by saying $\mathcal{V}^B = T(\mathcal{V}^B) \subset L^2(\Omega_T)$ and, analogously, we consider $v \in L^\infty(\Omega_T)$, if v is in $L^\infty(\mathcal{K}_\alpha^h)$ for all $\alpha \in \mathcal{A}$. The norms for $L^2(\Omega_T)$ and $L^\infty(\Omega_T)$ are denoted in the usual way by $\|\cdot\|_{L^2(\Omega_T)}$ and $\|\cdot\|_{L^\infty(\Omega_T)}$, respectively. The volume of any set $S \subset \mathbb{R}^n$ for $n \in \mathbb{N}$ is denoted by $|S|$, when it is defined.

Let $\gamma_\alpha : \mathcal{K}_\alpha^h \rightarrow \partial\mathcal{K}_\alpha^h$ denote the trace operator on each cell. Given a function $v \in \mathcal{V}^B$, two trace functions $v_{\alpha-}$ and $v_{\alpha+}$ are associated with v at the interface between two adjacent cells \mathcal{K}_α^h and \mathcal{K}_β^h . These functions are defined by setting

$$v_{\alpha-}|_{\partial\mathcal{K}_\alpha^h \cap \partial\mathcal{K}_\beta^h} = \gamma_\alpha(v|_{\mathcal{K}_\alpha^h}), \text{ and } v_{\alpha+}|_{\partial\mathcal{K}_\alpha^h \cap \partial\mathcal{K}_\beta^h} = \gamma_\beta(v|_{\mathcal{K}_\beta^h}).$$

Note the convention of defining $v_{\alpha-}$ as the trace derived for a function v that is continuous on the closure of each cell by taking the limit of v approaching the boundary of cell α from the interior, whereas $v_{\alpha+}$ would be the limit derived if approaching the boundary from the exterior. If $v \in H^1(\Omega_T) \cap \mathcal{V}^B$ the notation is suppressed, and we say $\gamma_\alpha(v|_{\mathcal{K}_\alpha^h}) = v|_{\partial\mathcal{K}_\alpha^h}$. The jump in trace values across a cell interface relative to the cell \mathcal{K}_α^h is denoted by

$$[v]_\alpha = v_{\alpha+} - v_{\alpha-}. \quad (2.4)$$

Let $\mathcal{K}_\alpha^{h,T} = \mathcal{K}_\alpha^h \times (0, T)$ and $\Gamma_\alpha^{h,T} = \partial\mathcal{K}_\alpha^h \times (0, T)$, to maintain a compact notation.

2.2. Definition of an adjoint problem for the error. Let \hat{n}_α be the unit, outward-pointing, normal vector on the boundary of each cell \mathcal{K}_α^h . Recall that by assumption, the solution u of (1.1) satisfies $u \in H^1(\Omega_T)$. Then multiplying (1.1) by $v \in \mathcal{V}^B$ (subject to periodic boundary conditions in space) and integrating by parts over $\mathcal{K}_\alpha^{h,T} \times (0, T)$ gives

$$\int_{\mathcal{K}_\alpha^{h,T}} ((\partial_t u) v - u \nabla \cdot (\vec{a}v)) dx dt + \int_{\Gamma_\alpha^{h,T}} (\vec{a}u|_{\partial\mathcal{K}_\alpha^h}) \cdot \hat{n}_\alpha v_{\alpha-} d\sigma dt = \int_{\mathcal{K}_\alpha^{h,T}} F v dx dt,$$

where σ is the usual measure on cell boundaries. The flux across the cell interfaces must be well defined on \mathcal{V}^B . The flux $f(u) = \vec{a}u$ is replaced with some rule based on the left and right values at the cell interface, denoted by $\hat{f}(v_{\alpha-}, v_{\alpha+})$, for $\alpha \in \mathcal{A}$ and all $v \in \mathcal{V}^B$. The rule must be consistent, so that for the true solution $u \in H^1(\Omega_T)$,

$$\hat{f}(u_{\alpha-}, u_{\alpha+}) = f(u) = \vec{a}u|_{\partial\mathcal{K}_\alpha^h}, \forall \alpha \in \mathcal{A}. \quad (2.5)$$

Inserting this flux rule above and summing over $\alpha \in \mathcal{A}$,

$$A(u, v) = \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} F v dx dt, \quad \forall v \in \mathcal{V}^B, \quad (2.6)$$

where

$$A(u, v) = \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} ((\partial_t u) v - u \nabla \cdot (\vec{a}v)) dx dt + \sum_{\alpha \in \mathcal{A}} \int_{\Gamma_\alpha^{h,T}} \hat{f}(u_{\alpha-}, u_{\alpha+}) \cdot \hat{n}_\alpha v_{\alpha-} d\sigma dt. \quad (2.7)$$

This method of extracting a bilinear form consistent with the strong formulation, in the sense of (2.6), for sufficiently smooth solutions could also be derived by developing a local DG method as described by Cockburn *et.al.* [9, 10] and enriching the trial and test spaces. If the flux function \hat{f} is nonlinear, a linearization must be performed about an approximation $\tilde{u} \in \mathcal{V}^B$ of the solution to (1.1) in order to derive the adjoint problem for the error in the approximation of the QoI. The adjoint problem then becomes

dependent on the structure of the linearized problem. There are reports where similar derivations of adjoint error equations include this complication and where the well-posedness of the resulting adjoint problems is not known, but is assumed, (e.g. [26]).

It is critical to understand at this point that no particular choice of \hat{f} has been made, and so we consider a convenient choice to eliminate technical issues associated with linearization. In the case of linear advection, the following upwind rule is an intuitive choice that is both consistent and linear:

$$\hat{f}(v_{\alpha-}, v_{\alpha+}) = \begin{cases} \bar{a}v_{\alpha-}, & \text{if } \bar{a} \cdot \hat{n}_\alpha > 0, \\ \bar{a}v_{\alpha+}, & \text{if } \bar{a} \cdot \hat{n}_\alpha < 0, \end{cases} \quad (2.8)$$

for all $\alpha \in \mathcal{A}$ and for all $v \in \mathcal{V}^B$. The case $\bar{a} \cdot \hat{n}_\alpha = 0$ may be ignored, without loss of generality, since it implies zero flux.

To derive the adjoint problem for the error, let $\tilde{u} \in \mathcal{V}^B$ be some approximation for u subject to periodic boundary conditions in space, and define $\eta = u - \tilde{u}$. Since \hat{f} depends on the sign of $\bar{a} \cdot \hat{n}_\alpha$, define $\mathcal{G}_\alpha^+(\bar{z}) = \{x \in \Gamma_\alpha^{h,T} : \bar{z} \cdot \hat{n}_\alpha > 0\}$ and $\mathcal{G}_\alpha^-(\bar{z}) = \{x \in \Gamma_\alpha^{h,T} : \bar{z} \cdot \hat{n}_\alpha < 0\}$, for any $\bar{z} \in \mathbb{R}^D$. It follows from (2.7)-(2.8) that

$$\begin{aligned} A(u, v) - A(\tilde{u}, v) &= A(\eta, v) = \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} ((\partial_t \eta) v - \eta \nabla \cdot (\bar{a} v)) \, dx \, dt \\ &\quad + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{G}_\alpha^+(\bar{a})} (\bar{a} \cdot \hat{n}_\alpha) \eta_{\alpha-} v_{\alpha-} \, d\sigma \, dt + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{G}_\alpha^-(\bar{a})} (\bar{a} \cdot \hat{n}_\alpha) \eta_{\alpha+} v_{\alpha-} \, d\sigma \, dt. \end{aligned} \quad (2.9)$$

Integrate by parts the first integral to obtain

$$\begin{aligned} A(\eta, v) &= \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} (-\partial_t v) \eta + v \nabla \cdot (\bar{a} \eta) \, dx \, dt + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^h} (v \eta)|_{t=T} \, dx \\ &\quad - \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^h} (v \eta)|_{t=0} \, dx - \sum_{\alpha \in \mathcal{A}} \int_{\Gamma_\alpha^{h,T}} (\bar{a} \cdot \hat{n}_\alpha) \eta_{\alpha-} v_{\alpha-} \, d\sigma \, dt \\ &\quad + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{G}_\alpha^+(\bar{a})} (\bar{a} \cdot \hat{n}_\alpha) \eta_{\alpha-} v_{\alpha-} \, d\sigma \, dt + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{G}_\alpha^-(\bar{a})} (\bar{a} \cdot \hat{n}_\alpha) \eta_{\alpha+} v_{\alpha-} \, d\sigma \, dt. \end{aligned} \quad (2.10)$$

Some cancellation occurs between the last two integrals above, and so (2.10) can be rewritten as

$$\begin{aligned} A(\eta, v) &= \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} (-\partial_t v) \eta + v \nabla \cdot (\bar{a} \eta) \, dx \, dt + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^h} (v \eta)|_{t=T} \, dx \\ &\quad - \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^h} (v \eta)|_{t=0} \, dx + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{G}_\alpha^-(\bar{a})} [\eta]_\alpha (\bar{a} \cdot \hat{n}_\alpha) v_{\alpha-} \, d\sigma \, dt. \end{aligned} \quad (2.11)$$

The integrals evaluated at the initial and final times may be dropped at this point to derive a dual bilinear form, which will be denoted by $A^*(v, \eta)$. These terms will simply reappear in the error equation in the relationship between A and A^* . The dual bilinear form is

$$\begin{aligned} A^*(v, \eta) &= \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} (-\partial_t v) \eta + v \nabla \cdot (\bar{a} \eta) \, dx \, dt \\ &\quad + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{G}_\alpha^-(\bar{a})} [\eta]_\alpha (\bar{a} \cdot \hat{n}_\alpha) v_{\alpha-} \, d\sigma \, dt. \end{aligned} \quad (2.12)$$

This dual bilinear form may be associated with the strong form of a certain reverse-in-time linear advection problem in the following way. Notice that on any boundary of a cell $\mathcal{K}_\alpha^{h,T}$, if $\vec{a} \cdot \hat{n}_\alpha < 0$ then there is an adjacent cell, say $\mathcal{K}_\beta^{h,T}$, such that $\hat{n}_\beta = -\hat{n}_\alpha$ and $\vec{a} \cdot \hat{n}_\beta > 0$. At this interface, $\eta_{\alpha+} = \eta_{\beta-}$ and $v_{\alpha-} = v_{\beta+}$. Inserting these relationships into (2.12), it follows

$$\begin{aligned} A^*(v, \eta) &= \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} (-\partial_t v) \eta + v \nabla \cdot (\vec{a} \eta) \, dx \, dt \\ &\quad - \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{G}_\alpha^-(\vec{a})} (\vec{a} \cdot \hat{n}_\alpha) \eta_{\alpha-} v_{\alpha-} d\sigma \, dt - \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{G}_\alpha^+(\vec{a})} (\vec{a} \cdot \hat{n}_\alpha) \eta_{\alpha-} v_{\alpha+} d\sigma \, dt. \end{aligned} \quad (2.13)$$

Equivalently,

$$\begin{aligned} A^*(v, \eta) &= \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} (-\partial_t v) \eta + v \nabla \cdot (\vec{a} \eta) \, dx \, dt \\ &\quad + \sum_{\alpha \in \mathcal{A}} \int_{\Gamma_\alpha^{h,T}} \hat{g}(v_{\alpha-}, v_{\alpha+}) \cdot \hat{n}_\alpha \eta_{\alpha-} d\sigma \, dt, \end{aligned} \quad (2.14)$$

where \hat{g} is a the dual upwind flux function defined by

$$\hat{g}(v_{\alpha-}, v_{\alpha+}) = \begin{cases} -\vec{a} v_{\alpha-}, & \text{if } \vec{a} \cdot \hat{n}_\alpha < 0, \\ -\vec{a} v_{\alpha+}, & \text{if } \vec{a} \cdot \hat{n}_\alpha > 0. \end{cases} \quad (2.15)$$

The dual bilinear form may be associated with the strong linear advection problem in one of two ways, depending on the QoI $\mathcal{M}(u)$. In case the QoI takes the form (1.3), the adjoint problem for the error is to find ϕ satisfying

$$\begin{aligned} -\partial_t \phi - \nabla \cdot (\vec{a} \phi) &= \psi, \quad \text{in } \Omega_T, \\ \phi(x, t = T) &= 0 \quad \text{in } \Omega, \end{aligned} \quad (2.16)$$

subject to periodic boundary conditions in space. In case the QoI takes the form (1.4), the adjoint problem for the error is to find ϕ satisfying

$$\begin{aligned} -\partial_t \phi - \nabla \cdot (\vec{a} \phi) &= 0, \quad \text{in } \Omega_T, \\ \phi(x, t = T) &= \psi(x) \quad \text{in } \Omega, \end{aligned} \quad (2.17)$$

also subject to periodic boundary conditions in space. If $\phi \in H^1(\Omega_T) \cap \mathcal{V}^B$, solves (2.16), then

$$A^*(\phi, \eta) = \mathcal{M}(\eta), \quad \forall \eta \in \{\eta \in \mathcal{V}^B + \text{periodic BCs}\}, \quad (2.18)$$

or if $\phi \in H^1(\Omega_T) \cap \mathcal{V}^B$ solves (2.17), then

$$A^*(\phi, \eta) = 0, \quad \forall \eta \in \{\eta \in \mathcal{V}^B + \text{periodic BCs}\}. \quad (2.19)$$

The dual form consistency properties (2.18)-(2.19) are precisely what is needed to derive a formula for the error. There could be choices other than (2.8) for \hat{f} since the goal is only to derive the formula for the error, though if one wanted to define the adjoint problem in a variational form using A^* rather than relying on the solvability of the strong form, then more restrictions would likely be placed on choices for \hat{f} . Here, it is not necessary to define the adjoint problem in a variational form; this is left for possible future work. The focus is instead on investigating the asymptotic properties of the computable error estimate derived in the following section, using a continuum reconstruction of the finite volume data independent of the underlying type of finite volume solver.

2.3. The error representation formula. The error representation formula may now be derived in the standard way. The presentation of the remainder of this report is simplified by discussing only the case when the QoI is of the form (1.4). This does not affect any of the subsequent analysis or the conclusions.

Denote by $e_{\mathcal{M}}(\tilde{u})$ the error in the QoI, that is,

$$e_{\mathcal{M}}(\tilde{u}) = \mathcal{M}(u) - \mathcal{M}(\tilde{u}) = \int_{\Omega} \psi(x) (u - \tilde{u}) dx. \quad (2.20)$$

The error formula is obtained by using the fact that $\phi(x, T) = \psi(x)$ in (2.19) and by inserting $\eta = u - \tilde{u}$ into (2.11)-(2.12) to derive the following relationship between A and A^* :

$$\begin{aligned} e_{\mathcal{M}}(\tilde{u}) &= \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_{\alpha}^h} ((u - \tilde{u}) \phi)|_{t=T} dx = A((u - \tilde{u}), \phi) - A^*(\phi, (u - \tilde{u})) \\ &\quad + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_{\alpha}^h} ((u - \tilde{u}) \phi)|_{t=0} dx \\ \Rightarrow e_{\mathcal{M}}(\tilde{u}) &= A((u - \tilde{u}), \phi) + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_{\alpha}^h} (u_0 - \tilde{u}|_{t=0}) \phi|_{t=0} dx. \end{aligned} \quad (2.21)$$

The exact adjoint solution can be approximated by $\tilde{\phi} \in \mathcal{V}^B$ using any suitable numerical method. A computable error estimate is then extracted by adding and subtracting $\tilde{\phi}$, yielding

$$\begin{aligned} e_{\mathcal{M}}(\tilde{u}) &= A((u - \tilde{u}), \phi - \tilde{\phi}) + A((u - \tilde{u}), \tilde{\phi}) + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_{\alpha}^h} (\phi - \tilde{\phi})|_{t=0} (u_0 - \tilde{u}|_{t=0}) dx \\ &\quad + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_{\alpha}^h} \tilde{\phi}|_{t=0} (u_0 - \tilde{u}|_{t=0}) dx. \end{aligned} \quad (2.22)$$

By virtue of (2.6),

$$\begin{aligned} e_{\mathcal{M}}(\tilde{u}) &= \tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi}) + A((u - \tilde{u}), \phi - \tilde{\phi}) \\ &\quad + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_{\alpha}^h} (\phi - \tilde{\phi})|_{t=0} (u_0 - \tilde{u}|_{t=0}) dx, \end{aligned} \quad (2.23)$$

where $\tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})$ is the computable error estimate:

$$\tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi}) = \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_{\alpha}^h} F \tilde{\phi} dx dt - A(\tilde{u}, \tilde{\phi}) + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_{\alpha}^h} \tilde{\phi}|_{t=0} (u_0 - \tilde{u}|_{t=0}) dx. \quad (2.24)$$

The exact error representation formula (2.23) states that the difference between the true error in the approximation of the QoI and the approximation $\tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})$ of this error is equal to a bilinear product of the error fields $u - \tilde{u}$ and $\phi - \tilde{\phi}$. An important implication for uncertainty quantification is that the accuracy of the error measurement $\tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})$ depends on the space-time correlation of these error fields. For example, it will be highly desirable to find $\tilde{\phi}$ that minimizes the error $\phi - \tilde{\phi}$ at points where $u - \tilde{u}$ is large.

It is not clear precisely how $e_{\mathcal{M}}(\tilde{u}) - \tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})$ will compare to the size of $e_{\mathcal{M}}(\tilde{u})$ in general if the same computational grid and similar numerical methods are used for

both the forward and adjoint solutions. If the error in the forward solution is large on one cell, the error in the adjoint approximation on that cell must be correspondingly smaller to compensate if high accuracy in the error estimate is desired. Therefore, in practice, it may be necessary to use a different grid or higher-order numerical method for the adjoint problem. In any case, it is clear that with a sufficiently well-resolved approximation of the adjoint solution, the error in the QoI may, in principle, be approximated to any desired precision. Since in this context the adjoint problem has the same structure as the forward problem, we will consider approximations of the adjoint solution by finite volume methods. Then the approximations \tilde{u} and $\tilde{\phi}$ will be defined using standard tensor-grid polynomial reconstruction and interpolation in time. It follows that some asymptotic properties of the computable error estimate may be predicted.

3. Reconstruction mappings for finite volume data. The approximations \tilde{u} and $\tilde{\phi}$ could be constructed in many ways. In this report, a map is constructed from a finite volume data set \bar{w} to a space-time approximation $\tilde{w} \in \mathcal{V}^B$ of a smooth function $U(x, t)$. This map is then used to define \tilde{u} and $\tilde{\phi}$ from two finite volume data sets on potentially different grids. First, cell average-preserving polynomials are defined at each discrete time $\{t^0 = 0, t^1, \dots, t^N = T\}$ on each finite volume cell. Then a piecewise interpolant of the data in time is used to define \tilde{w} .

A common method of data reconstruction in space is used (see e.g. [11]) that yields piecewise polynomials that are continuous on each cell, with jumps across cell interfaces. The reconstruction will be defined using a spatial grid \mathcal{K}^H , defined as in Section 2.1, but possibly distinct from the grid \mathcal{K}^h used to define the broken space \mathcal{V}^B . The spatial reconstruction of the finite volume data \bar{w} at time t^n on the cell \mathcal{K}_α^H is denoted by $\tilde{w}_\alpha^n(x)$. The construction of $\tilde{w}_\alpha^n(x)$ is achieved using the finite volume data on a tensor product stencil of cells containing \mathcal{K}_α^H , denoted by S_α . Since periodic boundary conditions are considered and it is assumed that the function U is smooth, an approximation of order H^{p+1} may be obtained by considering the stencil

$$S_\alpha = \{\beta \in \mathbb{Z}^D : \alpha_d - r \leq \beta_d \leq \alpha_d - r + p, d = 1, \dots, D\}, \quad (3.1)$$

for a fixed stencil shift r and by defining a tensor product polynomial of degree at most p in each coordinate of the form

$$\tilde{w}_\alpha^n(x) = \sum_{d=1}^D \sum_{i_d=0}^p c_{(i_1, \dots, i_D)} \prod_{d=1}^D (x_d)^{i_d}, \quad (3.2)$$

where the constants $c_{(i_1, \dots, i_D)}$ depend on n, β, r, p and \bar{w} . In the presence of non-periodic boundary conditions, the stencil can be shifted independently for each cell toward the interior of the domain. Define $\tilde{w}_\alpha^n(x)$ to be the unique polynomial of the form (3.2) satisfying

$$\frac{1}{|\mathcal{K}_\beta^H|} \int_{\mathcal{K}_\beta^H} \tilde{w}_\alpha^n(x) dx = \bar{w}_\beta^n, \quad \forall \beta \in S_\alpha. \quad (3.3)$$

The spatial reconstruction over Ω at time t^n is denoted by $\tilde{w}^n(x)$, where $\tilde{w}^n(x)|_{\mathcal{K}_\alpha^H} = \tilde{w}_\alpha^n(x)$.

Given the reconstructions $\tilde{w}^n(x)$ for $n = 0, 1, \dots, N$, the final continuum approximation $\tilde{w}(x, t)$ is constructed by interpolation in time. If $x \in \mathcal{K}_\alpha^H$ and $0 \leq n - r_n \leq$

$N - p - 1$, $\tilde{w}(x, t)$ is defined on $[t^n, t^{n+1}]$ as the unique polynomial interpolant of degree $p + 1$ satisfying

$$\tilde{w}(x, t^\mu) = \tilde{w}_\alpha^\mu(x), \quad \mu = n - r_n, n - r_n + 1, \dots, n - r_n + p + 1. \quad (3.4)$$

Away from the initial and final times, we set $r_n = r$. In case the interval $[t^n, t^{n+1}]$ is near the initial or final times, the value of r_n is adjusted accordingly. The interpolant in time is of degree $p + 1$ by necessity to achieve optimal approximation properties for the computable error estimate; this property will be clarified in Theorem 4.1.

The polynomial degree p can be chosen based on the accuracy of the finite volume method to ensure an optimal rate of convergence for smooth solutions. Since the computable error estimate requires evaluation of first derivatives, this means the first derivative in space and in time of the reconstruction should ideally converge at the same rate as the finite volume data, since in general the convergence rate of the finite volume data should limit the convergence rate of any approximations based on this data. The following result is the first step to proving rigorously the asymptotic properties of the computable error estimate and to justifying this statement.

We prove the result in two spatial dimensions. The generalization to more dimensions is straight-forward, though notationally quite cumbersome. The maximum time step size Δt is defined by

$$\Delta t = \max_{n=0,1,\dots,N-1} (\Delta t^{n+1} = t^{n+1} - t^n).$$

LEMMA 3.1. *Let $\{\bar{w}_\alpha^n\}$ be a finite volume data set approximating some function $U(x, t) \in C^\infty(\overline{\Omega_T})$ for $x \in \mathbb{R}^2$ with periodic boundary conditions in space, and denote the cell average errors by $\bar{e}_\alpha^n = \bar{U}_\alpha^n - \bar{w}_\alpha^n$, where \bar{U}_α^n satisfies*

$$\bar{U}_\alpha^n = \frac{1}{|\mathcal{K}_\alpha^H|} \int_{\mathcal{K}_\alpha^H} U(x, t^n) dx.$$

Assume a regular series of space-time grids as defined above with $H \leq \bar{H} \in \mathbb{R}^+$, such that there exists $\rho > 0$ satisfying

$$\rho H \leq \min_{\substack{\alpha_d=1,2,\dots,M_d \\ d=1,2}} (H_d)_{\alpha_d} \quad \text{and} \quad \rho \Delta t \leq \min_{n=0,1,\dots,N-1} \Delta t^{n+1}, \quad (3.5)$$

independent of N or M_d , and that there is a fixed $\lambda > 0$ such that

$$\Delta t \leq \lambda \min_{\substack{\alpha_d=1,2,\dots,M_d \\ d=1,2}} (H_d)_{\alpha_d}. \quad (3.6)$$

Further assume that the following approximation properties hold for constants $C_0, C_1 > 0$ and $C_2 > 0$ independent of N or M_d :

$$\max_{n=0,1,\dots,N} \max_{\alpha} |\bar{e}_\alpha^n| \leq C_0 H^s, \quad (3.7)$$

$$\max_{n=0,1,\dots,N} \max_{\alpha_1=1,2,\dots,M_1} |\bar{e}_{(\alpha_1,\alpha_2)}^n - \bar{e}_{(\alpha_1-1,\alpha_2)}^n| \leq C_1 H^{s+1}, \quad (3.8)$$

$$\max_{n=0,1,\dots,N} \max_{\alpha_2=1,2,\dots,M_2} |\bar{e}_{(\alpha_1,\alpha_2)}^n - \bar{e}_{(\alpha_1,\alpha_2-1)}^n| \leq C_1 H^{s+1}, \quad (3.9)$$

$$\text{and} \quad \max_{n=0,\dots,N-1} \max_{\substack{\alpha_d=1,\dots,M_d \\ d=1,2}} |\bar{e}_\alpha^{n+1} - \bar{e}_\alpha^n| \leq C_2 H^s \Delta t. \quad (3.10)$$

Let \tilde{w} be the space-time reconstruction of \bar{w} , and approximate U , as defined in (3.1)-(3.4), using piece-wise polynomials of degree p in space and a stencil shift index r . Then the following approximation properties hold:

$$\|U - \tilde{w}\|_{L^\infty(\Omega_T)} = \mathcal{O}(H^{p+1} + H^s), \quad (3.11)$$

$$\|\partial_{x_d}(U - \tilde{w})\|_{L^\infty(\Omega_T)} = \mathcal{O}(H^p + H^s), \quad d = 1, 2, \quad (3.12)$$

$$\max_{\substack{\alpha_d=1,\dots,M_d \\ d=1,2}} \|[U - \tilde{w}]_\alpha\|_{L^\infty(\Gamma_\alpha^{H,T})} = \mathcal{O}(H^{p+1} + H^{s+1}), \quad (3.13)$$

$$\text{and } \|\partial_t(U - \tilde{w})\|_{L^\infty(\Omega_T)} = \mathcal{O}(H^{p+1} + H^s), \quad (3.14)$$

where the jumps $[U - \tilde{w}]_\alpha$ are defined by (2.4).

REMARK 3.1. Given a consistent finite volume method, the assumptions (3.7)-(3.10) are usually valid for some $s > 0$. The particular value of s will depend upon the method and the problem. Regardless, once the value of s has been identified, the ensuing analysis holds.

Proof. Uniform bounds will be derived on an arbitrary space-time cell $\overline{\mathcal{K}_\alpha^H} \times [t^n, t^{n+1}]$ in 2D. At each time t^μ for $\mu = n - r_n, n - r_n + 1, \dots, n - r_n + p + 1$, $\tilde{U}_\alpha^\mu(x)$ is defined as the unique polynomial of degree at most p in each coordinate of the form

$$\tilde{U}_\alpha^\mu(x) = \sum_{i_1=0}^p \sum_{i_2=0}^p c_{(i_1, i_2)}(x_1)^{i_1} (x_2)^{i_2} \quad (3.15)$$

that preserves the cell averages of U on the stencil S_α (see (3.1) above). That is,

$$\frac{1}{|\mathcal{K}_\beta^H|} \int_{\mathcal{K}_\beta^H} \tilde{U}_\alpha^\mu(x) dx = \bar{U}_\beta^\mu, \quad \forall \beta \in S_\alpha. \quad (3.16)$$

Then $\tilde{U}_\alpha^\mu(x)$ is restricted to $\overline{\mathcal{K}_\alpha^H}$. Given any $x \in \overline{\mathcal{K}_\alpha^H}$, the function $\tilde{U}_\alpha(x, t)$ is then defined to be the unique polynomial of degree $p + 1$ in time interpolating $\tilde{U}_\alpha^\mu(x)$ for $\mu = n - r_n, n - r_n + 1, \dots, n - r_n + p + 1$. The necessary error bounds are then derived by decomposing the errors into components $U - \tilde{U}_\alpha$ and $\tilde{U}_\alpha - \tilde{w}$, then applying the triangle inequality:

$$|U - \tilde{w}| \leq |U - \tilde{U}_\alpha| + |\tilde{U}_\alpha - \tilde{w}|, \quad (3.17)$$

$$|\partial_{x_d}(U - \tilde{w})| \leq |\partial_{x_d}(U - \tilde{U}_\alpha)| + |\partial_{x_d}(\tilde{U}_\alpha - \tilde{w})|, \quad d = 1, 2, \quad (3.18)$$

$$|[U - \tilde{w}]_\alpha|_{\partial\mathcal{K}_\alpha^H} \leq |[U - \tilde{U}_\alpha]_\alpha|_{\partial\mathcal{K}_\alpha^H} + |[\tilde{U}_\alpha - \tilde{w}]_\alpha|_{\partial\mathcal{K}_\alpha^H}, \quad (3.19)$$

$$\text{and } |\partial_t(U - \tilde{w})| = |\partial_t(U - \tilde{U}_\alpha)| + |\partial_t(\tilde{U}_\alpha - \tilde{w})|. \quad (3.20)$$

The first errors occurring on the right hand side of (3.17)-(3.20) can be bounded by writing them as interpolation errors as follows. Let $V(x, t)$ be defined on Ω_T by

$$V(x, t) = \int_{(x_1)_{-r-3/2}}^{x_1} \int_{(x_2)_{-r-3/2}}^{x_2} U(x, t) dx.$$

This is just the *primitive function* concept used in the derivation of the PPM and ENO schemes, [12, 22, 23]. The integration domain is large enough to include all the

necessary grid points, taking into account reconstruction stencils, but is not infinite since then V would be infinite. Differentiation of the primitive function V recovers the function U ; $U(x, t) = \partial_{x_1} \partial_{x_2} V(x, t)$. The primitive function may be calculated in terms of cell averages of U . If $j_d \geq 1 - r$, $d = 1, 2$, then

$$V(x_{(j_1-1/2, j_2-1/2)}, t^\mu) = \sum_{d=1}^2 \sum_{\beta_d=-r}^{j_d-1} (H_1)_{\beta_1} (H_2)_{\beta_2} \bar{U}_\beta^\mu. \quad (3.21)$$

At each time t^μ for $\mu = n - r_n, n - r_n + 1, \dots, n - r_n + p + 1$, polynomials $\tilde{V}_\alpha^\mu(x)$ are defined on S_α as the unique interpolants satisfying $\tilde{V}_\alpha^\mu(x) = V(x, t^\mu)$ with $x = x_{(j_1-1/2, j_2-1/2)}$ for $j_d \in \{\alpha_d - r, \dots, \alpha_d - r + p + 1\}$, $d = 1, 2$. Define $\tilde{V}_\alpha(x, t)$ as the polynomial interpolant in time of degree $p + 1$ of the data $\{\tilde{V}_\alpha^\mu(x)\}$ for $\mu = n - r_n, n - r_n + 1, \dots, n - r_n + p + 1$. It is easily verified using (3.21) (e.g. [11]) that $\tilde{U}_\alpha(x, t) = \partial_{x_1} \partial_{x_2} \tilde{V}_\alpha(x, t)$ on $S_\alpha \times [t^{n-r_n}, t^{n-r_n+p+1}]$, and it holds that

$$U(x, t) - \tilde{U}_\alpha(x, t) = \partial_{x_1} \partial_{x_2} \left(V(x, t) - \tilde{V}_\alpha(x, t) \right)$$

on $S_\alpha \times [t^{n-r_n}, t^{n-r_n+p+1}]$. Applying standard interpolation results and the time step restriction (3.6), the first errors on the right of (3.17)-(3.20) satisfy uniform estimates of the form

$$\left. \begin{aligned} |U - \tilde{U}_\alpha| &= \mathcal{O}(H^{p+1}), & |\partial_{x_d}(U - \tilde{U}_\alpha)| &= \mathcal{O}(H^p), & d = 1, 2, \\ |[U - \tilde{U}_\alpha]_\alpha|_{\partial\mathcal{K}_\alpha^H} &= \mathcal{O}(H^{p+1}), & \text{and } |\partial_t(U - \tilde{U}_\alpha)| &= \mathcal{O}(H^{p+1}). \end{aligned} \right\} \quad (3.22)$$

The results (3.11)-(3.14) follow from (3.17)-(3.20) and (3.22) after showing that

$$\left. \begin{aligned} |\tilde{U}_\alpha - \tilde{w}| &= \mathcal{O}(H^s), & |\partial_{x_d}(\tilde{U}_\alpha - \tilde{w})| &= \mathcal{O}(H^s), & d = 1, 2, \\ |[\tilde{U}_\alpha - \tilde{w}]_\alpha|_{\partial\mathcal{K}_\alpha^H} &= \mathcal{O}(H^{s+1}), & \text{and } |\partial_t(\tilde{U}_\alpha - \tilde{w})| &= \mathcal{O}(H^s). \end{aligned} \right\} \quad (3.23)$$

These details are quite cumbersome in 2D and are presented for completeness in Appendices A-E. \square

4. Asymptotic properties of the computable error estimate. The rate at which the computable error estimate for a quantity of interest converges to the true error will be derived, given smooth data for the problems (1.1) and (2.17). The solutions $u(x, t)$ and $\phi(x, t)$ are approximated using the reconstructions $\tilde{u}(x, t)$ and $\tilde{\phi}(x, t)$, respectively, as defined in Section 3.

It may be desirable to compute the finite volume approximations used to define \tilde{u} and $\tilde{\phi}$ on different grids. The finite volume data sets are therefore not defined here relative to the grid \mathcal{K}^h . Consider a finite volume data set \bar{U} defined with respect to a potentially coarser grid \mathcal{K}^H with grid size $H \geq h$, requiring that \mathcal{K}^h is a refinement of \mathcal{K}^H . Let \tilde{w} be the continuum reconstruction for \bar{U} as defined in Section 3. Note that if \mathcal{V}^{B*} is the broken space defined relative to \mathcal{K}^H , then $\mathcal{V}^{B*} \subset \mathcal{V}^B$, hence $\tilde{w} \in \mathcal{V}^B$ as required. In other words, we may compute a finite volume approximation for u on one grid and define \tilde{u} , compute a finite volume approximation for ϕ on another grid and define $\tilde{\phi}$, and if \mathcal{K}^h is taken to be the union of the two grids then it will follow $\tilde{u} \in \mathcal{V}^B$ and $\tilde{\phi} \in \mathcal{V}^B$.

The reconstructed solution \tilde{u} consists of piece-wise polynomials of degree p_u that are based on finite volume data $\bar{v} = \{\bar{v}_\alpha^n\}$ defined relative to a grid \mathcal{K}^{H_u} , where \mathcal{K}^h is

a refinement of \mathcal{K}^{H_u} . The associated maximum time step size is Δt_u . The exact cell average data is $\{\bar{u}_\alpha^n\}$, and the errors in the finite volume data are denoted by

$$(\bar{e}_u)_\alpha^n = \bar{u}_\alpha^n - \bar{v}_\alpha^n. \quad (4.1)$$

Similarly, the adjoint solution ϕ is reconstructed using piece-wise polynomials of degree p_ϕ that are based on finite volume data $\bar{w} = \{\bar{w}_\beta^m\}$ defined relative to a grid \mathcal{K}^{H_ϕ} , where \mathcal{K}^h is also a refinement of \mathcal{K}^{H_ϕ} . The associated maximum time step size is Δt_ϕ . The exact cell average data is $\{\bar{\phi}_\beta^m\}$, and the errors in the finite volume data are denoted by

$$(\bar{e}_\phi)_\beta^m = \bar{\phi}_\beta^m - \bar{w}_\beta^m. \quad (4.2)$$

THEOREM 4.1. *Let the solution $u(x, t)$ of (1.1) and the solution $\phi(x, t)$ of (2.17) be in $\mathcal{C}^\infty(\bar{\Omega}_T)$. If the errors (4.1) satisfy (3.7) - (3.10) with $s = s_u \in \mathbb{R}^+$, $\Delta t = \Delta t_u$ and $H = H_u$, and if the errors (4.2) satisfy (3.7) - (3.10) with $s = s_\phi \in \mathbb{R}^+$, $\Delta t = \Delta t_\phi$ and $H = H_\phi$, then under the assumptions of Lemma 3.1 it follows that*

$$e_{\mathcal{M}}(\tilde{u}) = \tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi}) + \mathcal{O}(H_u^{q_u} H_\phi^{q_\phi}), \quad (4.3)$$

$$q_u \geq \min\{p_u + 1, s_u\}, \quad (4.4)$$

$$q_\phi \geq \min\{p_\phi, s_\phi\}, \quad (4.5)$$

where p_u and p_ϕ are defined in the two preceding paragraphs.

REMARK 4.1. *The relationships (4.4)-(4.5) are not presented as equalities to emphasize the possibility of superconvergent results that have been observed in practice. The key point is that a minimum order of accuracy for the computable error estimate may be expected under the assumptions of the theorem.*

Proof. The result is proved in two spatial dimensions since Lemma 3.1 is required, though both Lemma 3.1 and hence this theorem could be extended to three spatial dimensions; this is notationally cumbersome. It follows from (2.9) and (2.23) that

$$\begin{aligned} e_{\mathcal{M}}(\tilde{u}) - \tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi}) &= A((u - \tilde{u}), \phi - \tilde{\phi}) + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^h} (\phi - \tilde{\phi}) \Big|_{t=0} (u_0 - \tilde{u}|_{t=0}) dx \\ &= \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} (\partial_t(u - \tilde{u})) (\phi - \tilde{\phi}) dx dt - \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} (u - \tilde{u}) \nabla \cdot (\vec{a}(\phi - \tilde{\phi})) dx dt \\ &+ \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{G}_\alpha^+(\vec{a})} \vec{a} \cdot \hat{n}_\alpha (u - \tilde{u})_{\alpha-} (\phi - \tilde{\phi})_{\alpha-} d\sigma dt + \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{G}_\alpha^-(\vec{a})} \vec{a} \cdot \hat{n}_\alpha (u - \tilde{u})_{\alpha+} (\phi - \tilde{\phi})_{\alpha-} d\sigma dt \\ &+ \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^h} (\phi - \tilde{\phi}) \Big|_{t=0} (u_0 - \tilde{u}|_{t=0}) dx. \end{aligned} \quad (4.6)$$

Due to the periodic boundary conditions, the terms evaluated at cell interfaces can be regrouped as described in Section 2.2, so that

$$\begin{aligned} e_{\mathcal{M}}(\tilde{u}) - \tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi}) &= \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^h} (\phi - \tilde{\phi}) \Big|_{t=0} (u_0 - \tilde{u}|_{t=0}) dx \\ &+ \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} (\partial_t(u - \tilde{u})) (\phi - \tilde{\phi}) dx dt - \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} (u - \tilde{u}) \nabla \cdot (\vec{a}(\phi - \tilde{\phi})) dx dt \\ &+ \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{G}_\alpha^+(\vec{a})} \vec{a} \cdot \hat{n}_\alpha (u - \tilde{u})_{\alpha-} [\phi - \tilde{\phi}]_\alpha d\sigma dt. \end{aligned} \quad (4.7)$$

Proceeding term by term, (4.7) can be bounded using Lemma 3.1. The first term on the right hand side is bounded using (3.11) and (3.14):

$$\begin{aligned} \left| \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} \partial_t(u - \tilde{u})(\phi - \tilde{\phi}) dx dt \right| &\leq |\Omega_T| \|\partial_t(u - \tilde{u})\|_{L^\infty(\Omega_T)} \|\phi - \tilde{\phi}\|_{L^\infty(\Omega_T)} \\ &= \mathcal{O}\left((H_u^{p_u+1} + H_u^{s_u})(H_\phi^{p_\phi+1} + H_\phi^{s_\phi})\right). \end{aligned} \quad (4.8)$$

Similarly,

$$\left| \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^{h,T}} (u - \tilde{u}) \nabla \cdot (\vec{a}(\phi - \tilde{\phi})) dx dt \right| = \mathcal{O}\left((H_u^{p_u+1} + H_u^{s_u})(H_\phi^{p_\phi} + H_\phi^{s_\phi})\right),$$

by (3.11)-(3.12). The next term on the right hand side of (4.7) satisfies

$$\begin{aligned} \left| \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{G}_\alpha^+(\vec{a})} \vec{a} \cdot \hat{n}_\alpha (u - \tilde{u})_{\alpha-} [\phi - \tilde{\phi}]_\alpha d\sigma dt \right| \\ = \mathcal{O}\left(C(H_\phi) (H_u^{p_u+1} + H_u^{s_u})(H_\phi^{p_\phi+1} + H_\phi^{s_\phi+1})\right). \end{aligned} \quad (4.9)$$

by (3.11)-(3.13), where $C(H_\phi) = \max_{d=1,2} |a_d| T \{M_{\phi,1}(x_2^R - x_2^L) + M_{\phi,2}(x_1^R - x_1^L)\}$. Here $M_{\phi,d}$, $d = 1, 2$, is the number of cells across the spatial domain Ω in the coordinate direction x_d for the grid \mathcal{K}^{H_ϕ} . This corresponds to where $\tilde{\phi}$ may be discontinuous, since otherwise $[\phi - \tilde{\phi}]_\alpha = 0$. It is easily verified that

$$\{M_{\phi,1}(x_2^R - x_2^L) + M_{\phi,2}(x_1^R - x_1^L)\} \leq \frac{2|\Omega|}{\rho H_\phi}$$

so that

$$\begin{aligned} \left| \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{G}_\alpha^+(\vec{a})} \vec{a} \cdot \hat{n}_\alpha (u - \tilde{u})_{\alpha-} [\phi - \tilde{\phi}]_\alpha d\sigma dt \right| \\ = \mathcal{O}\left(\max_{d=1,2} |a_d| T \frac{2|\Omega|}{\rho H_\phi} (H_u^{p_u+1} + H_u^{s_u})(H_\phi^{p_\phi+1} + H_\phi^{s_\phi+1})\right) \\ = \mathcal{O}\left((H_u^{p_u+1} + H_u^{s_u})(H_\phi^{p_\phi} + H_\phi^{s_\phi})\right). \end{aligned} \quad (4.10)$$

The last term on the right hand side of (4.7) satisfies

$$\left| \sum_{\alpha \in \mathcal{A}} \int_{\mathcal{K}_\alpha^h} (u_0 - \tilde{u}|_{t=0})(\phi - \tilde{\phi})|_{t=0} dx \right| = \mathcal{O}\left((H_u^{p_u+1} + H_u^{s_u})(H_\phi^{p_\phi+1} + H_\phi^{s_\phi})\right). \quad (4.11)$$

by (3.11). It follows that

$$e_{\mathcal{M}}(\tilde{u}) - \tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi}) = \mathcal{O}\left(H_u^{p_u+1} H_\phi^{p_\phi} + H_u^{p_u+1} H_\phi^{s_\phi} + H_u^{s_u} H_\phi^{p_\phi} + H_u^{s_u} H_\phi^{s_\phi}\right),$$

from which the desired result is extracted. \square

5. Computational examples. Numerical tests for validation of the theory are presented that include various finite volume solvers, polynomial reconstructions, and quantities of interest. An example with a nonsmooth adjoint solution is included for reference, even though the derivation of the error representation formula does not apply directly in this case. In all the tests, (1.1) is unforced, i.e. $F(x, t) = 0$.

5.1. 1D test cases. The space-time domain is taken to be $[0, 1] \times [0, 1]$, and the advection coefficient is $a = 1$. The initial condition for the forward problem is

$$u_0(x) = \sin(2\pi(x - 1/4)). \quad (5.1)$$

The quantities of interest are defined as final time conditions for the adjoint problem (i.e. (2.17)), denoted by $\psi_j(x)$ for $j = 1, 2, 3$. Specifically,

$$\mathcal{M}(u) = \int_0^1 \psi_j(x) u(x, t = 1) dx,$$

where

$$\psi_1(x) = \exp(-i 2\pi x), \text{ (Fourier coefficient)} \quad (5.2)$$

$$\psi_2(x) = \delta(x - x^*), \text{ (evaluate at } x = x^*) \quad (5.3)$$

$$\psi_3(x) = \frac{1}{2} u(x, T). \text{ (solution energy)} \quad (5.4)$$

The centering of $\psi_2(x)$ about $x^* = (1 + \pi/16)/2 \in (1/2, 1)$ is chosen to avoid the collocation of x^* with any cell centers (where errors may be misleadingly small) or interfaces (where the reconstructed solution is not defined). The third quantity of interest is nonlinear with respect to the forward solution. We follow the standard approach [5, 6] and linearize about the reconstructed solution \tilde{u} . Normally one computes a Fréchet derivative as part of the linearization process, but for the solution energy, some algebraic arguments suffice to derive the approximation:

$$\begin{aligned} \mathcal{M}(u) - \mathcal{M}(\tilde{u}) &= \frac{1}{2} \int_0^1 (u|_{t=T})^2 dx - \frac{1}{2} \int_0^1 (\tilde{u}|_{t=T})^2 dx \\ &= \frac{1}{2} \int_0^1 (u + \tilde{u})|_{t=T} (u - \tilde{u})|_{t=T} dx \approx \int_0^1 \tilde{u}|_{t=T} (u - \tilde{u})|_{t=T} dx. \end{aligned} \quad (5.5)$$

The adjoint data $\psi_3(x)$ is approximated by $\bar{\psi}_3(x) = \tilde{u}(x, T)$. The additional error committed by this approximation is

$$(\mathcal{M}(u) - \mathcal{M}(\tilde{u})) - \int_0^1 \bar{\psi}_3(x) (u - \tilde{u})|_{t=T} dx = \frac{1}{2} \int_0^1 (u - \tilde{u})^2|_{t=T} dx. \quad (5.6)$$

The finite volume cells in space are defined relative to the same uniform grid \mathcal{K}^h for approximation of both the forward and adjoint solutions, and the same time steps are used. Three upwind finite volume methods are employed. The first solver is a first-order upwind scheme:

$$\bar{u}_j^{n+1} = \bar{u}_j^n - a \frac{\Delta t}{h} (\bar{u}_j^n - \bar{u}_{j-1}^n). \quad (5.7)$$

Note that the scheme is linear since a is constant. The second solver is the second-order upwind scheme of Fromm [20], implemented using unlimited piece-wise linear

reconstruction of the data on each cell \mathcal{K}_j^h at time t^n in the manner of a MUSCL scheme [33]:

$$\bar{u}_j^{n+1} = \bar{u}_j^n - a \frac{\Delta t}{h} \left(\bar{u}_j^n - \bar{u}_{j-1}^n + \frac{h - a\Delta t}{2} (\sigma_j^n - \sigma_{j-1}^n) \right), \quad (5.8)$$

$$\sigma_j^n = \frac{\bar{u}_{j+1}^n - \bar{u}_{j-1}^n}{2h}. \quad (5.9)$$

The third solver is a high-fidelity, total variation diminishing (TVD) MUSCL scheme that is the same as the second-order upwind scheme, except that the slopes are locally restricted using the minimum modulus limiter. In this case, the slopes are given by

$$\mu_j^n = \frac{1}{h} \text{MinMod}(\bar{u}_{j+1}^n - \bar{u}_j^n, \bar{u}_j^n - \bar{u}_{j-1}^n), \quad (5.10)$$

$$\text{MinMod}(\eta_1, \eta_2) = \begin{cases} \eta_1, & \text{if } |\eta_1| < |\eta_2| \text{ and } \eta_1 \eta_2 > 0, \\ \eta_2, & \text{if } |\eta_2| < |\eta_1| \text{ and } \eta_1 \eta_2 > 0, \\ 0, & \text{otherwise.} \end{cases} \quad (5.11)$$

Limiting the slopes in this way is a common procedure used to capture discontinuities without numerical oscillations.

We now define a system of labeling the numerical tests, which will be organized by QoI. Let the approximation \tilde{u} of the solution u of (1.1) be constructed by taking finite volume data generated by one of the above methods and performing a space-time reconstruction of the data as described in Section 3. Given one of the quantities of interest implied by (5.2)-(5.4), let $\tilde{\phi}$ be constructed in an analogous way as an approximation to the solution ϕ of (2.17). The stencil shift index is $r = 1$ (Section 3) in the tests, for both linear and quadratic reconstructions. There is an associated error in the quantity of interest, $e_{\mathcal{M}}(\tilde{u})$, and a computable estimate of this error, $\tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})$, as defined in (2.23)-(2.24). We denote by

$$\text{acc}(\tilde{u}, \tilde{\phi}) = e_{\mathcal{M}}(\tilde{u}) - \tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi}) \quad (5.12)$$

the accuracy of the error estimate. In each case there are four choices to be made:

- forward finite volume solver, denoted by FS*k* for $k = 1, 2, 3$;
- adjoint finite volume solver, denoted by AS*k* for $k = 1, 2, 3$; and
- spatial reconstructions (forward and adjoint), denoted by R*k* for $k = 1, 2$.

We used these codes to label each test. For example, the test label FS3-R2 AS1-R1 corresponds to choosing the high-fidelity forward solver with piece-wise quadratic spatial reconstruction, and the first-order upwind adjoint solver with piece-wise linear spatial reconstruction. Table 5.1 makes these definitions precise.

Each test consists of a convergence analysis for the quantities $\tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})$ and $\text{acc}(\tilde{u}, \tilde{\phi})$, with a fixed CFL number of $a\Delta t/h = 6/7$. In all the tests throughout Section 5 the convergence rates are computed as the slope of the line fitted by least-squares linear regression through the logarithm of the absolute values of the error data and grid sizes. Figures 5.1-5.6 plot the absolute values of $\tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})$ and $\text{acc}(\tilde{u}, \tilde{\phi})$. Solid black lines without markers are included in these plots that show the slopes associated with integer convergence rates, for reference.

5.1.1. 1D Fourier coefficient estimation. The results in this section pertain to the first non-constant Fourier coefficient quantity of interest, (5.2). This is an example where the solution of the adjoint problem is smooth, and the results of

TABLE 5.1
Convention for labeling finite volume solvers.

Label	FS1 or AS1	FS2 or AS2	FS3 or AS3
Solver	order 1 upwind	order 2 upwind	high fidelity (MinMod)

Theorem 4.1 apply directly. Since the finite volume approximations are defined on the same uniform grid we have $H_u = H_\phi = h$. The conclusion of Theorem 4.1 may be restated in the simplified form

$$e_{\mathcal{M}}(\tilde{u}) = \tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi}) + \mathcal{O}(h^{q_u + q_\phi}). \quad (5.13)$$

The computed convergence rates $q = q_u + q_\phi$ in (5.13), for the real and imaginary parts of the QoI, correspond to the two right-most columns in Table 5.2; the minimum values for q from Theorem 4.1 are listed alongside, in parentheses, for reference.

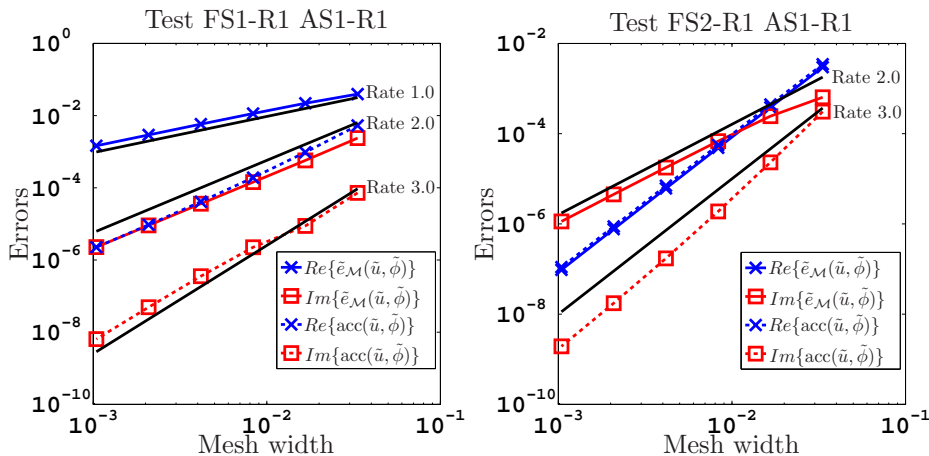


FIG. 5.1. Left: an example with first-order linear solvers. Right: A suboptimal reconstruction of the forward data is detrimental to the error estimate.

TABLE 5.2
Convergence rates for the tests in Section 5.1.1.

Test	Rate for $Re\{\tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})\}$	Rate for $Im\{\tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})\}$	Rate for $Re\{acc(\tilde{u}, \tilde{\phi})\}$	Rate for $Im\{acc(\tilde{u}, \tilde{\phi})\}$
FS1-R1 AS1-R1	0.96	2.00	2.24 (2.0)	2.64 (2.0)
FS2-R1 AS1-R1	2.99	1.85	2.99 (2.0)	3.46 (2.0)
FS2-R2 AS1-R1	3.00	1.95	3.96 (3.0)	3.35 (3.0)
FS2-R2 AS2-R2	3.00	1.99	5.06 (4.0)	5.26 (4.0)
FS3-R2 AS1-R1	2.04	1.87	3.04 (2.3)	3.32 (2.3)
FS3-R2 AS2-R2	2.05	1.92	5.17 (3.3)	4.14 (3.3)

It is expected that $q \geq 2$ with the first-order upwind finite volume method for the forward and adjoint problems. This is consistent with the results in Table 5.2 for the accuracy of the real and imaginary parts of the error estimate. A superconvergence result is observed for the imaginary part of the error. Though the imaginary part of the error is estimated reasonably well, this is not predicted (again, we can only predict

$q \geq 2$ but need $q > 2$) since superconvergence results for $e_{\mathcal{M}}$ need not translate to similar results for $\text{acc}(\tilde{u}, \tilde{\phi})$.

An example where $e_{\mathcal{M}}$ converges at a superoptimal rate but $\text{acc}(\tilde{u}, \tilde{\phi})$ does not is provided by test FS2-R1 AS1-R1, where the forward solution is estimated using the second-order upwind method and piece-wise linear reconstruction in space, (Figure 5.1). In this case, Theorem 4.1 predicts $q \geq 3$, so although the theory is not contradicted, an important limitation of the theory is illuminated. The accuracy of the real part of the computable error estimate is not very good, in the sense that it is highly desirable for $\text{Re}\{\text{acc}(\tilde{u}, \tilde{\phi})\}$ to be smaller than the error $\text{Re}\{e_{\mathcal{M}}(\tilde{u})\}$. Predicting this phenomenon would require an estimate of the form

$$\left| e_{\mathcal{M}}(\tilde{u}) - \tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi}) \right| \leq C \Delta x^q |e_{\mathcal{M}}(\tilde{u})|. \quad (5.14)$$

In the *a posteriori* literature for elliptic-type problems, there is a measurement of quality, ϵ , for the computable error estimate known as the *effectivity index*, (see e.g., [1]), which is the size of the ratio of the computable error estimate to the true error: $\epsilon = \tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})/e_{\mathcal{M}}(\tilde{u})$. It is desirable to be able to say $\epsilon \rightarrow 1$ asymptotically as $h \rightarrow 0$. Clearly, this would follow from the result (5.14). Such a result is not currently known to exist in the literature for hyperbolic conservation laws. However, test FS2-R1 AS1-R1 does not represent a cause for concern in practice, since the reconstruction of the forward solution chosen is suboptimal; it makes little sense to compute the finite volume approximation to second-order and then only reconstruct to first-order. This is improved in test FS2-R2 AS1-R1, as seen in Table 5.2. In the tests herein (as well as numerous other tests not reported), the only cases where the effectivity index was not observed to converge to unity were when $p_u < s_u$ or $p_\phi < s_\phi$.

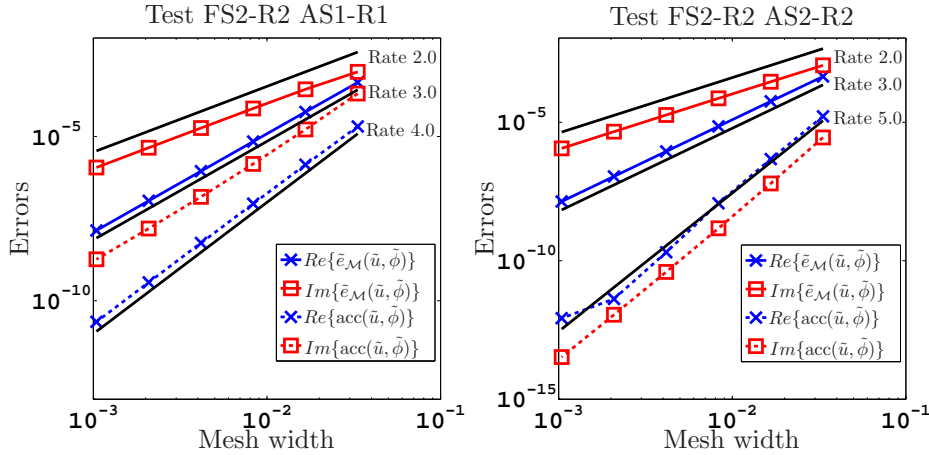


FIG. 5.2. Convergence tests with a second-order forward solver.

The tests in Figures 5.2-5.3 serve as examples of how the accuracy of the computable error estimate can be improved using higher-order finite volume methods for the adjoint problem. Conversely, if a highly accurate method is used for the forward problem, the error can still be estimated with some success using only a first-order solver for the adjoint problem. The adjoint error estimation technique is also effective when a nonlinear solver is used for the forward problem, e.g. Figure 5.3.

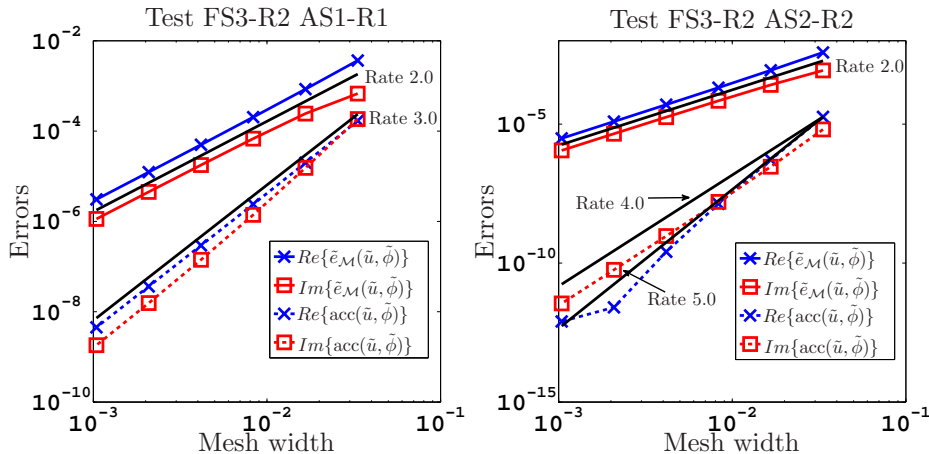


FIG. 5.3. Convergence tests with a nonlinear forward solver.

We note that there is a decrease in the convergence rate of $Re\{acc(\bar{u}, \bar{\phi})\}$ at the finest grid size for tests FS2-R2 AS2-R2 and FS3-R2 AS2-R2. At the larger grid sizes, the convergence rates of this quantity are unexpectedly high ($\approx 5 - 6$), and it is not uncommon in this situation to see the rate decline on finer grids due to the limits of machine precision. Also, the predicted values of q in Table 5.2 are lower for the tests using the TVD forward solver instead of the second-order upwind solver. This is because the convergence rate of the cell averages was measured as $s_u \approx 4/3$, in the sense of (3.7), for tests FS3-R2 AS1-R1 and FS3-R2 AS2-R2. The convergence rate is 2 for both $Re\{e_{\mathcal{M}}(\bar{u})\}$ and $Im\{e_{\mathcal{M}}(\bar{u})\}$. This is not surprising since the method is second-order accurate except on a finite number of cells where the true solution has extrema. Superconvergence was never observed for the true error using the TVD method.

5.1.2. Evaluation of the solution at a point in 1D. The quantity of interest is defined by (5.3). In this case, the adjoint data is a distribution, and Theorem 4.1 does not strictly apply. The goal is to demonstrate the practical limitations of using the proposed computable error estimate.

Convergence rates for the tests in this section are in Table 5.3. The absolute values of the computable error estimates and of the associated accuracy are plotted in Figure 5.4. The erratic convergence behavior of the error using the nonlinear forward solver is a known phenomenon, caused by the fact that the Min-Mod limiter invalidates Taylor's theorem. This was noted by Roache [30] to be a complication for applying Richardson extrapolation. The accuracy of the computable error estimate does not exhibit a clear convergence behavior for any of the tests in Figure 5.4. In this case, the accuracy values pass through zero during the sequence of grid refinements, causing the unpredictable behavior.

Though the relative scaling of the errors between consecutive grids is not predictable by the theory, it was observed that the error and the accuracy of the computable error estimate tend to be smaller using the higher-order finite volume methods. In all the tests, it was observed that the computable error estimate matched the true error in size, and usually the first 1-2 digits were correct. Also, most of the tests we have performed (including many not reported herein) have shown that the relative accuracy generally improves as the grid is refined for this quantity of interest.

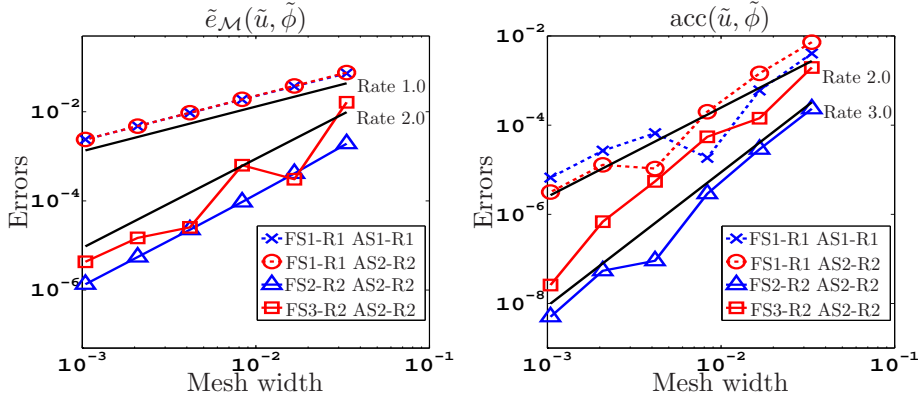


FIG. 5.4. Left: computable error estimate values for the solution value at a point. Right: accuracy of the computable error estimates.

Table 5.3 also includes the ranges of effectivity indices ϵ and the values of ϵ on the finest grid. The values of ϵ help to understand the quality of the error estimates without relying on asymptotic properties like convergence rates. The values of ϵ vary the most for test FS3-R2 AS2-R2. However, on the finest grid ϵ is much closer to unity in all cases than what is indicated by the ranges for ϵ , which indicates the general improvement of the error estimates upon grid refinement.

TABLE 5.3
Convergence rates for the tests in Section 5.1.2.

	FS1-R1 AS1-R1	FS1-R1 AS2-R2	FS2-R2 AS2-R2	FS3-R2 AS2-R2
Rates: $\tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})$	0.99	1.00	2.09	2.20
Rates: $\text{acc}(\tilde{u}, \tilde{\phi})$	1.65	2.30	3.13	3.07
ϵ : range	0.993 – 1.059	0.997 – 1.106	0.990 – 1.138	0.823 – 1.879
ϵ : fine grid	0.997	0.999	0.996	1.006

5.1.3. Final time solution energy in 1D. The linearization (5.5) is applied for the solution energy functional, committing an additional error in the estimation of the solution energy, shown in (5.6). The order of accuracy $q = q_u + q_\phi$ in (5.13) for the computable error estimate is thus amended:

$$q \geq \min \{p_u + p_\phi + 1, p_u + s_\phi + 1, s_u + p_\phi, s_u + s_\phi, 2s_u, 2p_u + 2\}. \quad (5.15)$$

The value on the right of this inequality is listed in Table 5.4, in parentheses, alongside the computed rate q .

TABLE 5.4
Convergence rates for the tests in Section 5.1.3.

	FS1-R1 AS1-R1	FS1-R1 AS2-R2	FS3-R2 AS1-R1
Rates: $\tilde{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})$	0.93	0.96	2.03
Rates: $\text{acc}(\tilde{u}, \tilde{\phi})$	2.12 (2.0)	1.98 (2.0)	3.10 (2.3)

In Table 5.4, the convergence rates for the accuracy of the computable error estimates are consistent with (5.15). Due to the linearization error, the accuracy

is limited to rate 2 for test FS1-R1 AS2-R2. Since the linearization error scales like order $2 \min\{s_u, p_u + 1\}$, this limitation is much less severe when a higher-order method is used for the forward problem. Indeed, in Table 5.4, the results using the TVD solver for the forward problem with a first-order upwind adjoint solver show that the accuracy increases to the expected nominal order of 3.

5.2. 2D Fourier coefficient estimation. The space-time domain is $\Omega \times [0, T] = [0, 1]^2 \times [0, 0.5]$, and the advection vector is $\vec{a} = \langle 0.5, 0.65 \rangle$. A uniform grid is used for each test with the total number of finite volume cells at each time step and the number of time steps both listed in the results. The ratio of time step size to cell size is $\Delta t / \Delta x = 5/9$. The initial data for the forward advection problem is

$$u_0(x_1, x_2) = \sin(2\pi(x_1 - 0.25)) \cos(2\pi(x_2 - 0.25)). \quad (5.16)$$

The quantity of interest is again a Fourier coefficient. In this case, the quantity of interest takes the form

$$\begin{aligned} \mathcal{M}(u) &= \int_{\Omega} \psi(x) u(x, T) dx, \quad \text{where} \\ \psi(x) &= \exp(-2\pi i k \cdot x), \quad k = \langle 1, 1 \rangle. \end{aligned}$$

Two finite volume methods are considered, the first being a first-order accurate method, derived by approximating the solution as constant on each cell and solving the corresponding Riemann problems in each coordinate direction. The second method is derived by a method of lines approach. The spatial discretization is performed by applying a second-order upwind flux approximation in each coordinate direction (this is just the 2D extension of (5.8)-(5.9)). The time stepping is performed using the classical explicit, fourth-order Runge-Kutta method. The same labeling scheme as was employed for the 1D tests is reused here, except *FS1* and *AS1* now refer to the two-dimensional, first-order method, while *FS2* and *AS2* refer to the second-order method of lines discretization. Also, *R1* and *R2* refer to the 2D reconstruction by a tensor product of linear polynomials, and of quadratic polynomials, respectively.

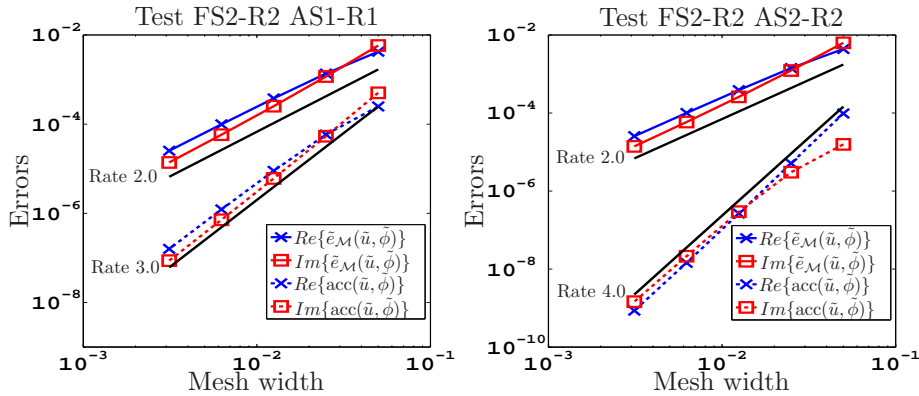


FIG. 5.5. Errors for 2D tests FS2-R2 AS1-R1 and FS2-R2 AS2-R2.

TABLE 5.5
Convergence rates for the tests in Section 5.2.

Test	Rate for $Re\{\bar{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})\}$	Rate for $Im\{\bar{e}_{\mathcal{M}}(\tilde{u}, \tilde{\phi})\}$	Rate for $Re\{acc(\tilde{u}, \tilde{\phi})\}$	Rate for $Im\{acc(\tilde{u}, \tilde{\phi})\}$
FS1-R1 AS1-R1	0.88	0.92	2.27 (2.0)	2.35 (2.0)
FS1-R1 AS1-R1-coarse	0.82	0.85	2.09 (2.0)	2.11 (2.0)
FS1-R1 AS2-R2	0.96	0.97	3.05 (3.0)	3.31 (3.0)
FS2-R2 AS1-R1	1.85	2.18	2.94 (3.0)	3.12 (3.0)
FS2-R2 AS2-R2	1.87	2.20	4.20 (4.0)	3.87 (4.0)

In every test with the exception of one, the same grid is used to estimate the solutions of the forward and adjoint problems. The forward and adjoint solutions are smooth, and the results of Theorem 4.1 apply directly. The results of the tests, varying the finite volume solvers used, are shown in Table 5.5 and Figures 5.5-5.6. The convergence rates of $Re\{acc(\tilde{u}, \tilde{\phi})\}$ in test FS2-R2 AS1-R1 and $Im\{acc(\tilde{u}, \tilde{\phi})\}$ in test FS2-R2 AS2-R2 are computed using only the two finest grids; in Figure 5.5 it is clear that at larger grid sizes this data is not in the asymptotic regime.

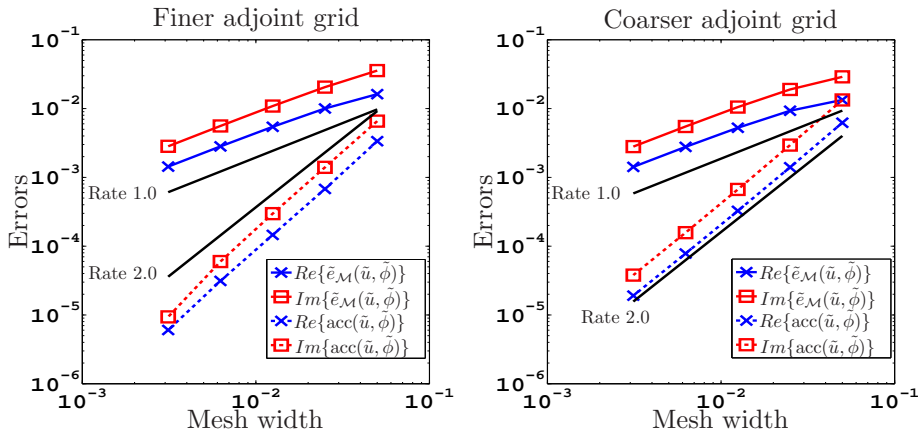


FIG. 5.6. Errors for 2D test FS1-R1 AS1-R1. Left: using the same grid for the forward and adjoint approximations. Right: using a coarser grid for the adjoint approximation.

One test is performed that is the same as test FS1-R1 AS1-R1, with the distinction that the finite volume approximation for the adjoint problem is now coarser than that used for the approximation of the forward solution. The same sequence of grids is used for the forward approximation, but the grids for the adjoint approximation contain cells that are a factor of 2 wider in space with time steps twice as large. This corresponds to a factor of eight reduction in the cost of running the adjoint solver. The convergence results are listed in Table 5.5 with the test label “FS1-R1 AS1-R1-coarse”, and the data is plotted in Figure 5.6. The expected behavior is confirmed: using the coarser grid for the adjoint approximation results in some loss of accuracy of the computable error estimate compared to the results using the finer grid, but the convergence rate of $acc(\tilde{u}, \tilde{\phi})$ is 2, which is consistent with Theorem 4.1.

No superconvergence results were observed in 2D. It is likely that, for many applications in multiple dimensions using nonlinear schemes, there would also not be a superconvergence of the true error. Thus it is promising that a first-order adjoint

solver with piece-wise linear reconstruction would suffice to estimate the error such that the effectivity index would converge asymptotically to 1.

6. Conclusions and future work. A method of computing *a posteriori* estimates of the error in a quantity of interest has been defined in the case of constant-coefficient linear advection with periodic boundary conditions. This method allows any suitable numerical discretization to be used to estimate the solution of the adjoint problem. A statement of the asymptotic behavior of the computable error estimate has been rigorously proven in Theorem 4.1 when the solutions to the forward and adjoint problems are smooth.

Several tests have been supplied that support the claims of Theorem 4.1. The results also show the theory to be pessimistic in some cases. Since the goal is to estimate the error in quantities of interest reliably, there are only two cases to consider. The first case is when the true error in a quantity of interest converges at the nominal rate associated with the forward solver method. In this case only a first-order accurate adjoint solver is required to guarantee that the *effectivity index* (Section 5.1.1) converges to 1. This is desirable in the absence of a computable upper bound for the error. The second case is due to the fact that superconvergence is sometimes observed for the true error. This needs to be accounted for when estimating the adjoint variable, since the error cancellation does not necessarily carry through to improve the accuracy of the computable error estimate. Therefore a higher-order adjoint solver may be needed to guarantee that the effectivity index converges to unity in this second case. However, superconvergence has not been observed in any of our tests in multiple dimensions nor when a nonlinear scheme has been applied for the forward problem.

A natural next step in further developing the proposed error estimation technique is the extension to the problem of a nonlinear hyperbolic conservation law with non-periodic boundary conditions. The method cannot currently be considered robust in the presence of shocks, though usually reasonable error estimates can be obtained. Also, it would be useful to characterize precisely when the ratio of the computable error estimate to the true error converges to 1, which remains an open problem.

Acknowledgments. The authors would like to thank Professor Donald Estep of Colorado State University for valuable discussions on adjoint methodology and for editorial feedback.

REFERENCES

- [1] MARK AINSWORTH AND J. TINSLEY ODEN, *A posteriori error estimation in finite element analysis*, Comput. Method Appl. M., 142 (1997), pp. 1–88.
- [2] KENDALL E. ATKINSON, *An Introduction to Numerical Analysis*, John Wiley and Sons, second ed., 1988.
- [3] GARRETT E. BARTER, *Shock capturing with PDE-based artificial viscosity for an adaptive, higher-order discontinuous Galerkin finite element method*, PhD thesis, MIT, 2008.
- [4] TIMOTHY J. BARTH, *A posteriori error estimation and mesh adaptivity for finite volume and finite element methods*, in Adaptive Mesh Refinement - Theory and Applications, Tomasz Plewa, Timur Linde, and V. Gregory Weirs, eds., vol. 41 of Lecture Notes in Computational Science and Engineering, Springer Berlin Heidelberg, 2005, pp. 183–202.
- [5] ———, *Space-time error representation and estimation in Navier-Stokes calculations*, in Complex Effects in Large Eddy Simulations, Stavros C. Kassinos, Carlos A. Langer, Gianluca Iaccarino, and Parviz Moin, eds., vol. 56 of Lecture Notes in Computational Science and Engineering, Springer Berlin Heidelberg, 2007, pp. 29–48.
- [6] TIMOTHY J. BARTH AND MATS G. LARSON, *A posteriori error estimates for higher order Godunov finite volume methods on unstructured meshes*, in Finite Volumes for Complex Applications III, R. Herbin and D. Kroner, HERMES Science Publishing Ltd, 2003.

- [7] OLAF BENEDIX AND BORIS VEXLER, *A posteriori error estimation and adaptivity for elliptic optimal control problems with state constraints*, *Comput. Optim. Appl.*, 44 (2009), pp. 3–25.
- [8] RICKARD E. BENSOW AND MATS G. LARSON, *Residual based VMS subgrid modeling for vortex flows*, *Comput. Method Appl. M.*, 199 (2010), pp. 802–809.
- [9] BERNARDO COCKBURN, SAN-YIH LIN, AND CHI-WANG SHU, *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws III: One-dimensional systems*, *J. Comput. Phys.*, 84 (1989), pp. 90 – 113.
- [10] BERNARDO COCKBURN AND CHI-WANG SHU, *The Runge-Kutta discontinuous Galerkin method for conservation laws V: Multidimensional systems*, *J. Comput. Phys.*, (1998), pp. 199 – 224.
- [11] BERNARDO COCKBURN, CHI-WANG SHU, CLAES JOHNSON, EITAN TADMOR, AND CHI-WANG SHU, *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*, in *Advanced Numerical Approximation of Nonlinear Hyperbolic Equations*, vol. 1697 of *Lecture Notes in Mathematics*, Springer Berlin / Heidelberg, 1998, pp. 325–432.
- [12] PHILLIP COLELLA AND PAUL R. WOODWARD, *The piecewise parabolic method (PPM) for gas-dynamical simulations*, *J. Comput. Phys.*, 54 (1984), pp. 174 – 201.
- [13] DON ESTEP, MATS G. LARSON, AND ROY D. WILLIAMS, *Estimating the Error of Numerical Solutions of Systems of Reaction-Diffusion Equations*, vol. 146 of *Memoirs of the American Mathematical Society*, AMS, Providence, RI, 2000.
- [14] DON ESTEP, MICHAEL PERNICE, DU PHAM, SIMON TAVENER, AND HAIYING WANG, *A posteriori error analysis of a cell-centered finite volume method for semilinear elliptic problems*, *J. Comput. Appl. Math.*, 233 (2009), pp. 459–472.
- [15] DON ESTEP, MICHAEL PERNICE, SIMON TAVENER, AND HAIYING WANG, *A posteriori error analysis for a cut cell finite volume method*, *Comput. Method Appl. M.*, 200 (2011), pp. 2768 – 2781. Special Issue on Modeling Error Estimation and Adaptive Modeling.
- [16] DON ESTEP, SIMON TAVENER, AND TIM WILDEY, *A posteriori error estimation and adaptive mesh refinement for a multiscale operator decomposition approach to fluid-solid heat transfer*, *J. Comput. Phys.*, 229 (2010), pp. 4143–4158.
- [17] KRZYSZTOF J. FIDKOWSKI AND DAVID L. DARMOFAL, *A triangular cut-cell adaptive method for high-order discretizations of the compressible Navier-Stokes equations*, *J. Comput. Phys.*, 225 (2007), pp. 1653 – 1672.
- [18] ———, *Output-based error estimation and mesh adaptation in computational fluid dynamics: overview and recent results*, in *Proc. of the 47th AIAA Aerospace Sciences Meeting*, 2009.
- [19] KRZYSZTOF J. FIDKOWSKI AND PHILIP L. ROE, *An entropy adjoint approach to mesh refinement*, *SIAM J. Sci. Comput.*, 32 (2010), pp. 1261–1287.
- [20] JACOB E. FROMM, *A method for reducing dispersion in convective difference schemes*, *J. Comput. Phys.*, 3 (1968), pp. 176–189.
- [21] MICHAEL B. GILES AND ENDRE SÜLI, *Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality*, *Acta Numerica*, 11 (2002), pp. 145–236.
- [22] AMI HARTEN, BJORN ENGQUIST, STANLEY OSHER, AND SUKUMAR R CHAKRAVARTHY, *Uniformly high order accurate essentially non-oscillatory schemes, III*, *J. Comput. Phys.*, 71 (1987), pp. 231 – 303.
- [23] AMI HARTEN, STANLEY OSHER, BJRN ENGQUIST, AND SUKUMAR R. CHAKRAVARTHY, *Some results on uniformly high-order accurate essentially nonoscillatory schemes*, *Appl. Numer. Math.*, 2 (1986), pp. 347 – 377. Special Issue in Honor of Milt Rose’s Sixtieth Birthday.
- [24] RALF HARTMANN, *Multitarget error estimation and adaptivity in aerodynamic flow simulations*, *SIAM J. Sci. Comput.*, 31 (2008), pp. 708–731.
- [25] RALF HARTMANN AND PAUL HOUSTON, *Adaptive discontinuous Galerkin finite element methods for nonlinear hyperbolic conservation laws*, *SIAM J. Sci. Comput.*, 24 (2002), pp. 979–1004.
- [26] ———, *Adaptive discontinuous Galerkin finite element methods for the compressible Euler equations*, *J. Comput. Phys.*, 183 (2002), pp. 508 – 532.
- [27] GURI I. MARCHUK, VALERI I. AGOSHKOV, AND VICTOR P. SHUTYAIEV, *Adjoint Equations and Perturbation Algorithms in Nonlinear Problems*, CRC Press, USA, 1996.
- [28] WILLIAM L OBERKAMPF, TIMOTHY G TRUCANO, AND CHARLES HIRSCH, *Verification, validation, and predictive capability in computational engineering and physics*, *Appl. Mech. Rev.*, 57 (2004), pp. 345–384.
- [29] NILES A. PIERCE AND MICHAEL B. GILES, *Adjoint and defect error bounding and correction for functional estimates*, *J. Comput. Phys.*, 200 (2004), pp. 769 – 794.
- [30] PATRICK J. ROACHE, *Verification and Validation in Computational Science and Engineering*, Hermosa publishers, New Mexico, USA, 1998.

- [31] CHRISTOPHER J. ROY, *Review of code and solution verification procedures for computational simulation*, J. Comput. Phys., 205 (2005), pp. 131 – 156.
- [32] CHRISTINA STEINER AND SEBASTIAN NOELLE, *On adaptive timestepping for weakly instationary solutions of hyperbolic conservation laws via adjoint error control*, Int. J. Numer. Meth. Biomed. Eng., 26 (2010), pp. 790–806.
- [33] BRAM VAN LEER, *Towards the ultimate conservative difference scheme. V. a second-order sequel to Godunov's method*, J. Comput. Phys., 32 (1979), pp. 101 – 136.

Appendix A. Proof of Lemma 3.1. The proof is continued by defining $\tilde{W}_\alpha(x, t)$ on $S_\alpha \times [t^{n-r_n}, t^{n-r_n+p+1}]$ to be the polynomial interpolant in space and time constructed the same way as $\bar{W}_\alpha(x, t)$, except that at grid vertices the approximate cell average data \bar{w}_β^μ replaces the exact cell average data:

$$\tilde{W}(x_{j_1-1/2, j_2-1/2}, t^\mu) = \sum_{d=1}^2 \sum_{\beta_d=-r}^{j_d-1} (H_1)_{\beta_1} (H_2)_{\beta_2} \bar{w}_\beta^\mu, \quad (\text{A.1})$$

with $j_d \geq 1 - r$ for $d = 1, 2$. Then just as it was shown that

$$U(x, t) - \tilde{U}_\alpha(x, t) = \partial_{x_1} \partial_{x_2} \left(V(x, t) - \tilde{V}_\alpha(x, t) \right)$$

on $S_\alpha \times [t^{n-r_n}, t^{n-r_n+p+1}]$, it also holds that

$$\tilde{U}_\alpha(x, t) - \tilde{w}_\alpha(x, t) = \partial_{x_1} \partial_{x_2} \left(\tilde{V}_\alpha(x, t) - \tilde{W}_\alpha(x, t) \right).$$

Given an arbitrary time level t^μ with $\mu \in \{n - r_n, \dots, n - r_n + p + 1\}$, define the polynomial $e_\alpha^\mu(x)$ as the difference $e_\alpha^\mu(x) = \tilde{V}_\alpha^\mu(x) - \tilde{W}_\alpha^\mu(x)$. Error bounds will first be derived at time t^μ on \mathcal{K}_α^H , then the extension of the error bounds in time will be described. The goal is to show that

$$\max_{x \in \mathcal{K}_\alpha^H} |\partial_{x_1} \partial_{x_2} e_\alpha^\mu(x)| = \mathcal{O}(H^s), \quad (\text{A.2})$$

$$\text{and } \max_{x \in \mathcal{K}_\alpha^H} |\partial_{x_d} \partial_{x_1} \partial_{x_2} e_\alpha^\mu(x)| = \mathcal{O}(H^s). \quad (\text{A.3})$$

Let $Z_d = \alpha_d - r$ for $d = 1, 2$. Then $e_\alpha^\mu(x)$ satisfies

$$e_\alpha^\mu(x) = \sum_{d=1}^2 \sum_{\beta_d=-r}^{j_d-1} (H_1)_{\beta_1} (H_2)_{\beta_2} \bar{e}_\beta^\mu, \quad (\text{A.4})$$

where $x = ((x_1)_{j_1-1/2}, (x_2)_{j_2-1/2})$ for $j_d = Z_d, \dots, Z_d + p + 1$ and $d = 1, 2$. Given an arbitrary $x \in \mathcal{K}_\alpha^H$, $e_\alpha^\mu(x)$ may be expanded using the Newton divided difference form of the interpolant (Atkinson [2]) in the coordinate x_1 , since it is a polynomial of degree $p + 1$ in x_1 interpolating the data $e_\alpha^\mu((x_1)_{j_1-1/2}, x_2)$:

$$\begin{aligned} e_\alpha^\mu(x) &= e_\alpha^\mu((x_1)_{Z_1-1/2}, x_2) + ((x_1) - (x_1)_{Z_1-1/2}) e_\alpha^\mu([(x_1)_{Z_1-1/2}, (x_1)_{Z_1+1/2}], x_2) \\ &\quad + ((x_1) - (x_1)_{Z_1-1/2})((x_1) - (x_1)_{Z_1+1/2}) e_\alpha^\mu([(x_1)_{Z_1-1/2}, \dots, (x_1)_{Z_1+3/2}], x_2) \\ &\quad + \dots + \prod_{m_1=0}^p ((x_1) - (x_1)_{Z_1+m_1-1/2}) e_\alpha^\mu([(x_1)_{Z_1-1/2}, \dots, (x_1)_{Z_1+p+1/2}], x_2). \end{aligned} \quad (\text{A.5})$$

The divided difference notation

$$e_\alpha^\mu([(x_1)_{j_1-1/2}, (x_1)_{j_1+1/2}, \dots, (x_1)_{l_1+j_1+1/2}], x_2)$$

for $l_1 = 0, 1, \dots, p$ and $j_1 = Z_1, \dots, Z_1 + p - l_1$ is defined first in case $l_1 = 0$, for $j_1 = Z_1, \dots, Z_1 + p$, by setting

$$e_\alpha^\mu \left([(x_1)_{j_1-1/2}, (x_1)_{j_1+1/2}], x_2 \right) = \frac{e_\alpha^\mu((x_1)_{j_1+1/2}, x_2) - e_\alpha^\mu((x_1)_{j_1-1/2}, x_2)}{(x_1)_{j_1+1/2} - (x_1)_{j_1-1/2}}$$

Then inductively with respect to l_1 , for $j_1 = Z_1, \dots, Z_1 + p - l_1$,

$$\begin{aligned} & e_\alpha^\mu \left([(x_1)_{j_1-1/2}, \dots, (x_1)_{j_1+l_1+1/2}], x_2 \right) \\ &= \frac{e_\alpha^\mu \left([(x_1)_{j_1+1/2}, \dots, (x_1)_{j_1+l_1+1/2}], x_2 \right)}{(x_1)_{j_1+l_1+1/2} - (x_1)_{j_1-1/2}} - \frac{e_\alpha^\mu \left([(x_1)_{j_1-1/2}, \dots, (x_1)_{j_1+l_1-1/2}], x_2 \right)}{(x_1)_{j_1+l_1+1/2} - (x_1)_{j_1-1/2}}. \end{aligned}$$

The mixed partial derivative $\partial_{x_1} \partial_{x_2} e_\alpha^\mu(x)$ must be bounded. Differentiate (A.5) to obtain the expression

$$\begin{aligned} \partial_{x_1} \partial_{x_2} e_\alpha^\mu(x) &= \partial_{x_2} e_\alpha^\mu \left([(x_1)_{Z_1-1/2}, (x_1)_{Z_1+1/2}], x_2 \right) \\ &+ \sum_{j_1=1}^p \sum_{l_1=0}^{j_1} \Psi(l_1, j_1, \alpha_1, x_1) \partial_{x_2} e_\alpha^\mu \left([(x_1)_{Z_1-1/2}, \dots, (x_1)_{Z_1+j_1+1/2}], x_2 \right), \end{aligned} \quad (\text{A.6})$$

where for $l_1 = 0, 1, \dots, j_1$ and $j_1 = 1, 2, \dots, p$,

$$\Psi(l_1, j_1, \alpha_1, x_1) = \prod_{\substack{m_1=0 \\ m_1 \neq l_1}}^{j_1} (x_1 - (x_1)_{Z_1+m_1-1/2}). \quad (\text{A.7})$$

The problem is reduced to bounding errors along grid lines as x_2 varies. The following result is used, proved in Appendix B:

Claim A1

There exists a fixed $C > 0$, independent of μ , H or α , such that if $(x_2)_{\alpha_2-1/2} \leq x_2 \leq (x_2)_{\alpha_2+1/2}$, then

$$\left| \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+1/2}], x_2 \right) \right| \leq C H^s, \quad (\text{A.8})$$

for $m_1 = Z_1, \dots, Z_1 + p$. It follows that

$$\begin{aligned} & \left| \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, \dots, (x_1)_{m_1+3/2}], x_2 \right) \right| \\ & \leq \left| \frac{\partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1+1/2}, (x_1)_{m_1+3/2}], x_2 \right)}{(x_1)_{m_1+3/2} - (x_1)_{m_1-1/2}} \right| + \left| \frac{\partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+1/2}], x_2 \right)}{(x_1)_{m_1+3/2} - (x_1)_{m_1-1/2}} \right| \\ & \leq \frac{2C H^s}{2\rho H} = \frac{C H^{s-1}}{\rho}, \end{aligned} \quad (\text{A.9})$$

for $m_1 = Z_1, \dots, Z_1 + p - 1$. Using induction on j_1 , it is shown that

$$\left| \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, \dots, (x_1)_{m_1+j_1+1/2}], x_2 \right) \right| \leq C \frac{2^{j_1} H^{s-j_1}}{(j_1 + 1)! \rho^{j_1}}, \quad (\text{A.10})$$

where $m_1 = Z_1, \dots, Z_1 + p - j_1$, for $j_1 = 1, \dots, p$. Case $j_1 = 1$ holds by (A.9). Assuming case j_1 holds, apply the definition of the divided difference to show case

$j_1 + 1$ holds (explicit details are provided in the proof of Claim A1). Then the desired uniform bound on $\partial_{x_1} \partial_{x_2} e_\alpha^\mu$ is derived by first bounding (A.7) as follows:

$$|\Psi(l_1, j_1, \alpha_1, x_1)| = \left| \prod_{\substack{m_1=0 \\ m_1 \neq l_1}}^{j_1} (x_1 - (x_1)_{Z_1+m_1-1/2}) \right| \leq ((p+1)H)^{j_1}. \quad (\text{A.11})$$

Inserting this result in (A.6) and applying (A.8) and (A.10) with $m_1 = Z_1$:

$$\begin{aligned} |\partial_{x_1} \partial_{x_2} e_\alpha^\mu(x)| &\leq C H^s + C \sum_{j_1=1}^p (j_1 + 1) ((p+1)H)^{j_1} \frac{2^{j_1} H^{s-j_1}}{(j_1 + 1)! \rho^{j_1}} \\ &= C \left(1 + \sum_{j_1=1}^p (j_1 + 1) (p+1)^{j_1} \frac{2^{j_1}}{(j_1 + 1)! \rho^{j_1}} \right) H^s. \end{aligned} \quad (\text{A.12})$$

This bound holds on any cell \mathcal{K}_α^H , proving (A.2). In proving the result (A.3), we may choose $d = 1$ without loss of generality, as the proof may be repeated in an analogous way for case $d = 2$. In (A.6)-(A.7), differentiate with respect to x_1 to obtain

$$\begin{aligned} \partial_{x_1}^{(2)} \partial_{x_2} e_\alpha^\mu(x) &= 2 \partial_{x_2} e_\alpha^\mu([(x_1)_{Z_1-1/2}, \dots, (x_1)_{Z_1+3/2}], x_2) \\ &+ \sum_{j_1=2}^p \sum_{l_1=0}^{j_1} (\partial_{x_1} \Psi(l_1, j_1, \alpha_1, x_1)) \partial_{x_2} e_\alpha^\mu([(x_1)_{Z_1-1/2}, \dots, (x_1)_{Z_1+j_1+1/2}], x_2), \end{aligned} \quad (\text{A.13})$$

where

$$\partial_{x_1} \Psi(l_1, j_1, \alpha_1, x_1) = \sum_{\substack{i_1=0 \\ i_1 \neq l_1}}^{j_1} \prod_{\substack{m_1=0 \\ m_1 \neq l_1, i_1}}^{j_1} (x_1 - (x_1)_{Z_1+m_1-1/2}) \quad (\text{A.14})$$

for $l_1 = 0, 1, \dots, j_1$ and $j_1 = 2, \dots, p$. Another claim is made, bounding the first term on the right hand side of (A.13) by a fixed positive constant, again denoted by $C > 0$. The proof is provided in Appendix C.

Claim A2

There exists a fixed $C > 0$, independent of μ, H or α , such that if $(x_2)_{\alpha_2-1/2} \leq x_2 \leq (x_2)_{\alpha_2+1/2}$, then for $m_1 = Z_1, \dots, Z_1 + p - 1$,

$$|\partial_{x_2} e_\alpha^\mu([(x_1)_{m_1-1/2}, \dots, (x_1)_{m_1+3/2}], x_2)| \leq C H^s. \quad (\text{A.15})$$

Proceeding as above, an inductive argument is used to show the following bound, derived using Claim A2:

$$|\partial_{x_2} e_\alpha^\mu([(x_1)_{m_1-1/2}, \dots, (x_1)_{m_1+j_1+1/2}], x_2)| \leq C \frac{2^{j_1} H^{s+1-j_1}}{(j_1 + 1)! \rho^{j_1-1}} \quad (\text{A.16})$$

for $m_1 = Z_1, \dots, Z_1 + p - j_1$ and $j_1 = 2, 3, \dots, p$. Bound (A.14) by

$$|\partial_{x_1} \Psi(l_1, j_1, \alpha_1, x_1)| \leq \sum_{\substack{i_1=0 \\ i_1 \neq l_1}}^{j_1} \prod_{\substack{m_1=0 \\ m_1 \neq l_1, i_1}}^{j_1} |x_1 - (x_1)_{Z_1+m_1-1/2}| \leq j_1 ((p+1)H)^{j_1-1}$$

and insert this result with (A.15) and (A.16) (with $m_1 = 0$) into (A.13), yielding

$$\left| \partial_{x_1}^{(2)} \partial_{x_2} e_{\alpha}^{\mu}(x) \right| \leq C H^s \left(2 + \sum_{j_1=2}^p \sum_{l_1=0}^{j_1} j_1 (p+1)^{j_1-1} \frac{2^{j_1}}{(j_1+1)! \rho^{j_1-1}} \right). \quad (\text{A.17})$$

This bound holds on every cell \mathcal{K}_{α}^H , proving (A.3).

Next, the jump of the error across a cell wall at time t^n is bounded. The arguments are shown here for the cell wall $x_1 = (x_1)_{\alpha_1-1/2}$ and $(x_2)_{\alpha_2-1/2} \leq x_2 \leq (x_2)_{\alpha_2+1/2}$, without loss of generality. The goal is to show that

$$\partial_{x_1} \partial_{x_2} e_{(\alpha_1+1, \alpha_2)}^{\mu}((x_1)_{\alpha_1-1/2}, x_2) - \partial_{x_1} \partial_{x_2} e_{(\alpha_1, \alpha_2)}^{\mu}((x_1)_{\alpha_1-1/2}, x_2) = \mathcal{O}(H^{s+1}). \quad (\text{A.18})$$

Applying (A.6)-(A.7), the difference (A.18) is

$$\begin{aligned} & \partial_{x_1} \partial_{x_2} e_{(\alpha_1+1, \alpha_2)}^{\mu}(x_1, x_2) - \partial_{x_1} \partial_{x_2} e_{(\alpha_1, \alpha_2)}^{\mu}(x_1, x_2) = \\ & \partial_{x_2} e_{(\alpha_1+1, \alpha_2)}^{\mu}([(x_1)_{Z_1+1/2}, (x_1)_{Z_1+3/2}], x_2) - \partial_{x_2} e_{(\alpha_1, \alpha_2)}^{\mu}([(x_1)_{Z_1-1/2}, (x_1)_{Z_1+1/2}], x_2) \\ & + \sum_{j_1=1}^p \sum_{l_1=0}^{j_1} \Psi(l_1, j_1, \alpha_1+1, x_1) \partial_{x_2} e_{(\alpha_1+1, \alpha_2)}^{\mu}([(x_1)_{Z_1+1/2}, \dots, (x_1)_{Z_1+j_1+3/2}], x_2) \\ & - \sum_{j_1=1}^p \sum_{l_1=0}^{j_1} \Psi(l_1, j_1, \alpha_1, x_1) \partial_{x_2} e_{(\alpha_1, \alpha_2)}^{\mu}([(x_1)_{Z_1-1/2}, \dots, (x_1)_{Z_1+j_1+1/2}], x_2) \quad (\text{A.19}) \end{aligned}$$

with $x_1 = (x_1)_{\alpha_1-1/2}$. Applying (A.15) the last two terms on the right hand side can be bounded, first by using

$$\begin{aligned} \left| \partial_{x_2} e_{(\alpha_1+1, \alpha_2)}^{\mu}([(x_1)_{Z_1+1/2}, \dots, (x_1)_{Z_1+1+3/2}], x_2) \right| & \leq C H^s, \\ \left| \partial_{x_2} e_{(\alpha_1, \alpha_2)}^{\mu}([(x_1)_{Z_1-1/2}, \dots, (x_1)_{Z_1+1+1/2}], x_2) \right| & \leq C H^s, \quad (\text{A.20}) \end{aligned}$$

and then proceeding inductively with respect to j_1 ,

$$\begin{aligned} \left| \partial_{x_2} e_{(\alpha_1+1, \alpha_2)}^{\mu}([(x_1)_{Z_1+1/2}, \dots, (x_1)_{Z_1+j_1+3/2}], x_2) \right| & \leq \frac{C H^s 2^{j_1}}{(j_1+1)! \rho^{j_1-1} H^{j_1-1}}, \\ \left| \partial_{x_2} e_{(\alpha_1, \alpha_2)}^{\mu}([(x_1)_{Z_1-1/2}, \dots, (x_1)_{Z_1+j_1+1/2}], x_2) \right| & \leq \frac{C H^s 2^{j_1}}{(j_1+1)! \rho^{j_1-1} H^{j_1-1}}. \quad (\text{A.21}) \end{aligned}$$

Insert (A.20)-(A.21) and (A.11) into (A.19) to obtain

$$\begin{aligned} & \left| \partial_{x_1} \partial_{x_2} e_{(\alpha_1+1, \alpha_2)}^{\mu}((x_1)_{\alpha_1+1/2}, x_2) - \partial_{x_1} \partial_{x_2} e_{(\alpha_1, \alpha_2)}^{\mu}((x_1)_{\alpha_1+1/2}, x_2) \right| \leq \\ & \left| \partial_{x_2} e_{(\alpha_1+1, \alpha_2)}^{\mu}([(x_1)_{Z_1+1/2}, (x_1)_{Z_1+3/2}], x_2) - \partial_{x_2} e_{(\alpha_1, \alpha_2)}^{\mu}([(x_1)_{Z_1-1/2}, (x_1)_{Z_1+1/2}], x_2) \right| \\ & + 2 \sum_{j_1=1}^p \sum_{l_1=0}^{j_1} ((p+1)H)^{j_1} C \frac{2^{j_1} H^{s+1-j_1}}{(j_1+1)! \rho^{j_1-1}}. \quad (\text{A.22}) \end{aligned}$$

The last term on the right hand side is of order H^{s+1} . The remaining terms are also of order H^{s+1} , which we list here as Claim A3, proved in Appendix D.

Claim A3

There exists a fixed $C > 0$, independent of μ , H or α , such that if $(x_2)_{\alpha_2-1/2} \leq x_2 \leq (x_2)_{\alpha_2+1/2}$, then

$$\left| \partial_{x_2} e_{(\alpha_1+1, \alpha_2)}^\mu \left([(x_1)_{Z_1+1/2}, (x_1)_{Z_1+3/2}], x_2 \right) - \partial_{x_2} e_{(\alpha_1, \alpha_2)}^\mu \left([(x_1)_{Z_1-1/2}, (x_1)_{Z_1+1/2}], x_2 \right) \right| \leq C H^{s+1}. \quad (\text{A.23})$$

Uniform bounds in space on cell \mathcal{K}_α^H have been shown at an arbitrary time t^μ . The error bounds must be extended to be uniform in time, and the time derivative of the error also needs to be bounded. Define $e_\alpha(x, t)$ on $\overline{\mathcal{K}_\alpha^H} \times [t^n, t^{n+1}]$ for each $x \in \overline{\mathcal{K}_\alpha^H}$ as the unique polynomial interpolant in time of degree $p+1$ such that

$$e_\alpha(x, t^\mu) = e_\alpha^\mu(x), \quad (\text{A.24})$$

for $\mu = n - r_n, \dots, n - r_n + p + 1$. The Newton form of the interpolant is

$$e_\alpha(x, t) = e_\alpha(x, t^{n-r_n}) + (t - t^{n-r_n}) e_\alpha(x, [t^{n-r_n}, t^{n-r_n+1}]) + \dots \quad (\text{A.25})$$

Proceeding as described above it can be shown that a uniform bound exists in terms of the interpolated data:

$$\max_{t \in [t^n, t^{n+1}]} \left| \partial_{x_d}^\omega \partial_{x_1} \partial_{x_2} e_\alpha(x, t) \right| \leq C \max_\mu \left| \partial_{x_d}^\omega \partial_{x_1} \partial_{x_2} e_\alpha^\mu(x) \right|, \quad (\text{A.26})$$

with C independent of n , α , or H , for $\omega = 0, 1$. Then the uniform bounds derived above are used for the right hand side, and an analogous result follows for the jump across $\partial \mathcal{K}_\alpha^H$. Taking derivatives in (A.25),

$$\partial_t \partial_{x_1} \partial_{x_2} e_\alpha(x, t) = \partial_{x_1} \partial_{x_2} e_\alpha(x, [t^{n-r_n}, t^{n-r_n+1}]) + \dots \quad (\text{A.27})$$

and to bound the time derivative of the error it suffices to show that:

Claim A4

$$\partial_{x_1} \partial_{x_2} e_\alpha(x, [t^{m_t}, t^{m_t+1}]) \leq C H^s, \quad (\text{A.28})$$

with C independent of n , α , or H , for $m_t = n - r_n, \dots, n - r_n + p$. This is shown in Appendix E. Then the remainder of the proof proceeds again as described above.

Appendix B. Proof of Claim A1.

Proof. Let x_1 satisfy $(x_1)_{Z_1-1/2} \leq x_1 \leq (x_1)_{Z_1+p+1/2}$. Then $\partial_{x_2} e_\alpha^\mu(x_1, x_2)$ may be expanded in x_2 using the Newton form of the interpolant:

$$\begin{aligned} \partial_{x_2} e_\alpha^\mu(x_1, x_2) &= \partial_{x_2} e_\alpha^\mu(x_1, (x_2)_{Z_2-1/2}) \\ &+ \partial_{x_2} \sum_{j_2=0}^p \prod_{l_2=0}^{j_2} (x_2 - (x_2)_{Z_2+l_2-1/2}) e_\alpha^\mu(x_1, [(x_2)_{Z_2-1/2}, \dots, (x_2)_{Z_2+j_2+1/2}]) \\ &= e_\alpha^\mu(x_1, [(x_2)_{Z_2-1/2}, (x_2)_{Z_2+1/2}]) \\ &+ \sum_{j_2=1}^p \sum_{q_2=0}^{j_2} \prod_{\substack{l_2=0 \\ l_2 \neq q_2}}^{j_2} (x_2 - (x_2)_{Z_2+l_2-1/2}) e_\alpha^\mu(x_1, [(x_2)_{Z_2-1/2}, \dots, (x_2)_{Z_2+j_2+1/2}]). \end{aligned} \quad (\text{B.1})$$

The divided difference that must be bounded is

$$\begin{aligned} & \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+1/2}], x_2 \right) \\ &= \frac{\partial_{x_2} e_\alpha^\mu \left((x_1)_{m_1+1/2}, x_2 \right) - \partial_{x_2} e_\alpha^\mu \left((x_1)_{m_1-1/2}, x_2 \right)}{(H_1)_{m_1}}. \end{aligned} \quad (\text{B.2})$$

for $m_1 = Z_1, \dots, Z_1 + p$. The numerator is expanded in x_2 by choosing $x_1 = (x_1)_{m_1+1/2}$ in (B.1), subtracting the analogous result with $x_1 = (x_1)_{m_1-1/2}$, and inserting the difference in (B.2). This expression is

$$\begin{aligned} & \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+1/2}], x_2 \right) = \\ & \frac{e_\alpha^\mu \left((x_1)_{m_1+1/2}, [(x_2)_{Z_2-1/2}, (x_2)_{Z_2+1/2}] \right) - e_\alpha^\mu \left((x_1)_{m_1-1/2}, [(x_2)_{Z_2-1/2}, (x_2)_{Z_2+1/2}] \right)}{(H_1)_{m_1}} \\ & + \sum_{j_2=1}^p \sum_{q_2=0}^{j_2} \frac{\Phi(q_2, j_2, \alpha_2)}{(H_1)_{m_1}} e_\alpha^\mu \left((x_1)_{m_1+1/2}, [(x_2)_{Z_2-1/2}, \dots, (x_2)_{Z_2+j_2+1/2}] \right) \\ & - \sum_{j_2=1}^p \sum_{q_2=0}^{j_2} \frac{\Phi(q_2, j_2, \alpha_2)}{(H_1)_{m_1}} e_\alpha^\mu \left((x_1)_{m_1-1/2}, [(x_2)_{Z_2-1/2}, \dots, (x_2)_{Z_2+j_2+1/2}] \right), \end{aligned} \quad (\text{B.3})$$

where $\Phi(q_2, j_2, \alpha_2) = \prod_{\substack{l_2=0 \\ l_2 \neq q_2}}^{j_2} (x_2 - (x_2)_{Z_2+l_2-1/2})$. The first two terms on the right hand side of (B.3) may be calculated in terms of the cell average errors, using the values of $e_\alpha^\mu(x)$ with $x = ((x_1)_{Z_1+j_1-1/2}, (x_2)_{Z_2+j_2-1/2})$, defined in (A.4):

$$e_\alpha^\mu(x) = \sum_{d=1}^2 \sum_{\beta_d=-r}^{Z_d+j_d-1} (H_1)_{\beta_1} (H_2)_{\beta_2} \bar{e}_\beta^\mu, \quad (\text{B.4})$$

for $j_d = 0, 1, \dots, p+1$ and $d = 1, 2$. It follows that

$$\begin{aligned} & \frac{e_\alpha^\mu \left((x_1)_{m_1+1/2}, [(x_2)_{Z_2-1/2}, (x_2)_{Z_2+1/2}] \right)}{(H_1)_{m_1}} \\ &= \frac{\sum_{\beta_1=-r}^{m_1} \sum_{\beta_2=-r}^{Z_2} (H_1)_{\beta_1} (H_2)_{\beta_2} \bar{e}_\beta^\mu - \sum_{\beta_1=-r}^{m_1} \sum_{\beta_2=-r}^{Z_2-1} (H_1)_{\beta_1} (H_2)_{\beta_2} \bar{e}_\beta^\mu}{(H_1)_{m_1} (H_2)_{Z_2}} \\ &= \frac{1}{(H_1)_{m_1}} \sum_{\beta_1=-r}^{m_1} (H_1)_{\beta_1} \bar{e}_{(\beta_1, Z_2)}^\mu, \end{aligned} \quad (\text{B.5})$$

and

$$\frac{e_\alpha^\mu \left((x_1)_{m_1-1/2}, [(x_2)_{Z_2-1/2}, (x_2)_{Z_2+1/2}] \right)}{(H_1)_{m_1}} = \frac{1}{(H_1)_{m_1}} \sum_{\beta_1=-r}^{m_1-1} (H_1)_{\beta_1} \bar{e}_{(\beta_1, Z_2)}^\mu. \quad (\text{B.6})$$

Therefore,

$$\begin{aligned} & \frac{e_\alpha^\mu \left((x_1)_{m_1+1/2}, [(x_2)_{Z_2-1/2}, (x_2)_{Z_2+1/2}] \right) - e_\alpha^\mu \left((x_1)_{m_1-1/2}, [(x_2)_{Z_2-1/2}, (x_2)_{Z_2+1/2}] \right)}{(H_1)_{m_1}} \\ &= \frac{1}{(H_1)_{m_1}} \left\{ \sum_{\beta_1=-r}^{m_1} (H_1)_{\beta_1} \bar{e}_{(\beta_1, Z_2)}^\mu - \sum_{\beta_1=-r}^{m_1-1} (H_1)_{\beta_1} \bar{e}_{(\beta_1, Z_2)}^\mu \right\} = \bar{e}_{(m_1, Z_2)}^\mu \end{aligned} \quad (\text{B.7})$$

for $m_1 = Z_1, \dots, Z_1 + p$. Similar arguments show that

$$\frac{e_\alpha^\mu \left((x_1)_{m_1+1/2}, [(x_2)_{m_2-1/2}, (x_2)_{m_2+1/2}] \right)}{(H_1)_{m_1}} - \frac{e_\alpha^\mu \left((x_1)_{m_1-1/2}, [(x_2)_{m_2-1/2}, (x_2)_{m_2+1/2}] \right)}{(H_1)_{m_1}} = \bar{e}_{(m_1, m_2)}^\mu \quad (\text{B.8})$$

for $m_1 = Z_1, \dots, Z_1 + p$ and $m_2 = Z_2, \dots, Z_2 + p$. An induction argument for j_2 is used to show the final result, i.e. bounding (B.3). Indeed, for case $j_2 = 1$,

$$\left| \frac{1}{(H_1)_{m_1}} e_\alpha^\mu \left((x_1)_{m_1+1/2}, [(x_2)_{m_2-1/2}, \dots, (x_2)_{m_2+j_2+1/2}] \right) - \frac{1}{(H_1)_{m_1}} e_\alpha^\mu \left((x_1)_{m_1-1/2}, [(x_2)_{m_2-1/2}, \dots, (x_2)_{m_2+j_2+1/2}] \right) \right| = \left| \frac{\bar{e}_{(m_1, m_2+1)}^\mu - \bar{e}_{(m_1, m_2)}^\mu}{(H_2)_{m_2+1} + (H_2)_{m_2}} \right| \leq \frac{C_1 H^{s+1}}{2\rho H} \quad (\text{B.9})$$

for $m_1 = Z_1, \dots, Z_1 + p$ and $m_2 = Z_2, \dots, Z_2 + p - j_2$, by assumption (3.9) of Lemma 3.1. Then expanding out the divided difference it is shown that case j_2 implies case $j_2 + 1$, and it holds that

$$\left| \frac{1}{(H_1)_{m_1}} e_\alpha^\mu \left((x_1)_{m_1+1/2}, [(x_2)_{m_2-1/2}, \dots, (x_2)_{m_2+j_2+1/2}] \right) - \frac{1}{(H_1)_{m_1}} e_\alpha^\mu \left((x_1)_{m_1-1/2}, [(x_2)_{m_2-1/2}, \dots, (x_2)_{m_2+j_2+1/2}] \right) \right| \leq C_1 \frac{2^{j_2} H^{s+1-j_2}}{(j_2 + 1)! \rho^{j_2}},$$

for $m_1 = Z_1, \dots, Z_1 + p$ and $m_2 = Z_2, \dots, Z_2 + p - j_2$, where $j_2 = 1, \dots, p$. Insert this result and (B.8) into (B.3), yielding the bound

$$\left| \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+1/2}], x_2 \right) \right| \leq C_1 \sum_{j_2=1}^p \sum_{q_2=0}^{j_2} |\Phi(q_2, j_2, \alpha_2)| \frac{2^{j_2} H^{s+1-j_2}}{(j_2 + 1)! \rho^{j_2}} + \left| \bar{e}_{(m_1, Z_2)}^\mu \right| \leq C_1 H^{s+1} \sum_{j_2=1}^p (j_2 + 1)(p + 1)^{j_2} \frac{2^{j_2}}{(j_2 + 1)! \rho^{j_2}} + C_0 H^s \quad (\text{B.10})$$

by the assumption (3.7) of Lemma 3.1. Recall from the statement of Lemma 3.1 that $H \leq \bar{H}$, hence $H^{s+1} \leq \bar{H} \cdot H^s$ and the desired result is achieved, choosing

$$C = C_0 + C_1 \bar{H} \sum_{j_2=1}^p (j_2 + 1)(p + 1)^{j_2} \frac{2^{j_2}}{(j_2 + 1)! \rho^{j_2}}.$$

□

Appendix C. Proof of Claim A2.

Proof. This result follows largely from the arguments in Appendix B. Expand the divided difference $\partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, \dots, (x_1)_{m_1+3/2}], x_2 \right)$ as follows:

$$\partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, \dots, (x_1)_{m_1+3/2}], x_2 \right) = \frac{\partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1+1/2}, (x_1)_{m_1+3/2}], x_2 \right)}{(H_1)_{m_1+1} + (H_1)_{m_1}} - \frac{\partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+1/2}], x_2 \right)}{(H_1)_{m_1+1} + (H_1)_{m_1}}. \quad (\text{C.1})$$

The goal is to show the numerator is uniformly of order H^{s+1} , and the desired result follows by applying the bound

$$\frac{1}{(H_1)_{m_1+1} + (H_1)_{m_1}} \leq \frac{1}{2\rho H}$$

for the denominator. Note that in (B.10) it is shown that all of the terms are uniformly of order H^{s+1} except for $\bar{e}_{(m_1, Z_2)}^\mu$, and in fact

$$\begin{aligned} \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1+1/2}, (x_1)_{m_1+3/2}], x_2 \right) &= \bar{e}_{(m_1+1, Z_2)}^\mu + \mathcal{O}(H^{s+1}) \\ \text{and } \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+1/2}], x_2 \right) &= \bar{e}_{(m_1, Z_2)}^\mu + \mathcal{O}(H^{s+1}). \end{aligned}$$

Subtracting,

$$\begin{aligned} \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1+1/2}, (x_1)_{m_1+3/2}], x_2 \right) - \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+1/2}], x_2 \right) \\ = \bar{e}_{(m_1+1, Z_2)}^\mu - \bar{e}_{(m_1, Z_2)}^\mu + \mathcal{O}(H^{s+1}) = \mathcal{O}(H^{s+1}), \end{aligned} \quad (\text{C.2})$$

by assumption (3.8) of Lemma 3.1. \square

Appendix D. Proof of Claim A3.

Proof. In this case the proof is almost the same as for Claim A2, using (B.10). Following the steps in the proof of Claim A1, it is shown that

$$\begin{aligned} \partial_{x_2} e_{(\alpha_1+1, \alpha_2)}^\mu \left([(x_1)_{Z_1+1/2}, (x_1)_{Z_1+3/2}], x_2 \right) &= \bar{e}_{(Z_1+1, Z_2)}^\mu + \mathcal{O}(H^{s+1}) \\ \text{and } \partial_{x_2} e_{(\alpha_1, \alpha_2)}^\mu \left([(x_1)_{Z_1-1/2}, (x_1)_{Z_1+1/2}], x_2 \right) &= \bar{e}_{(Z_1, Z_2)}^\mu + \mathcal{O}(H^{s+1}). \end{aligned}$$

Subtracting,

$$\begin{aligned} \partial_{x_2} e_{(\alpha_1+1, \alpha_2)}^\mu \left([(x_1)_{Z_1+1/2}, (x_1)_{Z_1+3/2}], x_2 \right) - \partial_{x_2} e_{(\alpha_1, \alpha_2)}^\mu \left([(x_1)_{Z_1-1/2}, (x_1)_{Z_1+1/2}], x_2 \right) \\ = \bar{e}_{(Z_1+1, Z_2)}^\mu - \bar{e}_{(Z_1, Z_2)}^\mu + \mathcal{O}(H^{s+1}) = \mathcal{O}(H^{s+1}), \end{aligned} \quad (\text{D.1})$$

by assumption (3.8) of Lemma 3.1. \square

Appendix E. Proof of Claim A4.

Proof. The divided difference to be bounded is

$$\partial_{x_1} \partial_{x_2} e_\alpha(x, [t^{m_t}, t^{m_t+1}]) = \frac{\partial_{x_1} \partial_{x_2} e_\alpha^{m_t+1}(x) - \partial_{x_1} \partial_{x_2} e_\alpha^{m_t}(x)}{\Delta t^{m_t+1}} \quad (\text{E.1})$$

for $m_t = n - r_n, \dots, n - r_n + p$. It suffices to show that

$$|\partial_{x_1} \partial_{x_2} e_\alpha^{m_t+1}(x) - \partial_{x_1} \partial_{x_2} e_\alpha^{m_t}(x)| \leq C H^s \Delta t, \quad (\text{E.2})$$

so that by (E.1) the final result follows: $|\partial_{x_1} \partial_{x_2} e_\alpha(x, [t^{m_t}, t^{m_t+1}])| \leq \frac{C H^s \Delta t}{\Delta t^{m_t+1}} \leq \frac{C H^s}{\rho}$.

The difference in (E.2) is rewritten via (A.6)-(A.7) from Lemma 3.1, so that

$$\begin{aligned} \partial_{x_1} \partial_{x_2} e_\alpha^{\mu+1}(x) - \partial_{x_1} \partial_{x_2} e_\alpha^\mu(x) &= \\ \partial_{x_2} e_\alpha^{\mu+1} \left([(x_1)_{Z_1-1/2}, (x_1)_{Z_1+1/2}], x_2 \right) - \partial_{x_2} e_\alpha^\mu \left([(x_1)_{Z_1-1/2}, (x_1)_{Z_1+1/2}], x_2 \right) \\ + \sum_{j_1=1}^p \sum_{l_1=0}^{j_1} \Psi(l_1, j_1, \alpha_1, x_1) \partial_{x_2} e_\alpha^{\mu+1} \left([(x_1)_{Z_1-1/2}, \dots, (x_1)_{Z_1+j_1+1/2}], x_2 \right) \\ - \sum_{j_1=1}^p \sum_{l_1=0}^{j_1} \Psi(l_1, j_1, \alpha_1, x_1) \partial_{x_2} e_\alpha^\mu \left([(x_1)_{Z_1-1/2}, \dots, (x_1)_{Z_1+j_1+1/2}], x_2 \right), \end{aligned} \quad (\text{E.3})$$

where $\Psi(l_1, j_1, \alpha_1, x_1) = \prod_{\substack{m_1=0 \\ m_1 \neq l_1}}^{j_1} (x_1 - (x_1)_{Z_1+m_1-1/2})$ for $l_1 = 0, 1, \dots, j_1$ and $j_1 = 1, 2, \dots, p$. Here $\mu = m_t$ for any arbitrary but fixed $m_t \in \{n - r_n, \dots, n - r_n + p\}$. The result (E.2) can be shown first by proving that

$$\left| \partial_{x_2} e_\alpha^{\mu+1} \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+1/2}], x_2 \right) - \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+1/2}], x_2 \right) \right| \leq C^* H^s \Delta t \quad (\text{E.4})$$

for any $m_1 = Z_1, \dots, Z_1 + p$, where $C^* > 0$ is independent of x_2 , α , μ and H . Then by induction, one proves (analogous to Appendix B-C) that

$$\left| \partial_{x_2} e_\alpha^{\mu+1} \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+j_1+1/2}], x_2 \right) - \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+j_1+1/2}], x_2 \right) \right| \leq C^* \frac{2^{j_1} H^{s-j_1} \Delta t}{(j_1 + 1)! \rho^{j_1}}, \quad (\text{E.5})$$

for $m_1 = Z_1, \dots, Z_1 + p - j_1$ and $j_1 = 1, 2, \dots, p$. Then (E.2) follows by applying (E.4)-(E.5) to (E.3) with the bound $|\Psi(l_1, j_1, \alpha_1, x_1)| \leq ((p+1)H)^{j_1}$. The inequality (E.4) is shown by using this result from Appendix B:

$$\begin{aligned} & \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+1/2}], x_2 \right) = \\ & \frac{e_\alpha^\mu \left((x_1)_{m_1+1/2}, [(x_2)_{Z_2-1/2}, (x_2)_{Z_2+1/2}] \right) - e_\alpha^\mu \left((x_1)_{m_1-1/2}, [(x_2)_{Z_2-1/2}, (x_2)_{Z_2+1/2}] \right)}{(H_1)_{m_1}} \\ & + \sum_{j_2=1}^p \sum_{q_2=0}^{j_2} \frac{\Phi(q_2, j_2, \alpha_2)}{(H_1)_{m_1}} e_\alpha^\mu \left((x_1)_{m_1+1/2}, [(x_2)_{Z_2-1/2}, \dots, (x_2)_{Z_2+j_2+1/2}] \right) \\ & - \sum_{j_2=1}^p \sum_{q_2=0}^{j_2} \frac{\Phi(q_2, j_2, \alpha_2)}{(H_1)_{m_1}} e_\alpha^\mu \left((x_1)_{m_1-1/2}, [(x_2)_{Z_2-1/2}, \dots, (x_2)_{Z_2+j_2+1/2}] \right). \quad (\text{E.6}) \end{aligned}$$

As described in Appendix B, the first two terms on the right hand side of (E.6) sum to be precisely $\bar{e}_{(m_1, Z_2)}^\mu$, and hence

$$\begin{aligned} & \partial_{x_2} e_\alpha^{\mu+1} \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+1/2}], x_2 \right) \\ & - \partial_{x_2} e_\alpha^\mu \left([(x_1)_{m_1-1/2}, (x_1)_{m_1+1/2}], x_2 \right) = \bar{e}_{(m_1, Z_2)}^{\mu+1} - \bar{e}_{(m_1, Z_2)}^\mu + \dots, \quad (\text{E.7}) \end{aligned}$$

where $\left| \bar{e}_{(m_1, Z_2)}^{\mu+1} - \bar{e}_{(m_1, Z_2)}^\mu \right| \leq C_2 H^s \Delta t$, by the assumption (3.10) of Lemma 3.1. Starting with this base case, the remaining terms in (E.7) are bounded using the usual induction arguments to prove (E.4) and hence the final result. This is tedious and follows previous arguments, so the remainder of the proof is omitted. \square