

FAST MULTIGRID REDUCTION-IN-TIME FOR ADVECTION VIA MODIFIED SEMI-LAGRANGIAN COARSE-GRID OPERATORS*

H. DE STERCK[†], R. D. FALGOUT[‡], AND O. A. KRZYSIK[§]

Abstract. Many iterative parallel-in-time algorithms have been shown to be highly efficient for diffusion-dominated partial differential equations (PDEs), but are inefficient or even divergent when applied to advection-dominated PDEs. We consider the application of the multigrid reduction-in-time (MGRIT) algorithm to linear advection PDEs. The key to efficient time integration with this method is using a coarse-grid operator that provides a sufficiently accurate approximation to the the so-called ideal coarse-grid operator. For certain classes of semi-Lagrangian discretizations, we present a novel semi-Lagrangian-based coarse-grid operator that leads to fast and scalable multilevel time integration of linear advection PDEs. The coarse-grid operator is composed of a semi-Lagrangian discretization followed by a correction term, with the correction designed so that the leading-order truncation error of the composite operator is approximately equal to that of the ideal coarse-grid operator. Parallel results show substantial speed-ups over sequential time integration for variable-wave-speed advection problems in one and two spatial dimensions, and using high-order discretizations up to order five. The proposed approach establishes the first practical method that provides small and scalable MGRIT iteration counts for advection problems.

Key words. parallel-in-time, MGRIT, Parareal, hyperbolic PDE, advection equation, multigrid

AMS subject classifications. 65F10, 65M22, 65M55, 35L03

1. Introduction. Traditionally, the solutions of initial-value partial differential equation (PDE) problems are approximated numerically via the sequential process of time-stepping, an approach motivated by the temporally causal nature of the solution itself. However, there also exist many time-parallel methods for simulating these PDEs, and these can often yield reductions in wall-clock time relative to time-stepping. A history and broad survey of the field of parallel-in-time methods can be found in the review [19], with the later review [32] providing a summary of more recent developments.

Presently, parallel-in-time methods are not widely used for large-scale PDE simulations, with the traditional technique of time-stepping remaining the standard. As supercomputer architectures continue to use more and more cores, parallel-in-time methods are likely to become essential for circumventing sequential time-stepping bottlenecks. However, one issue likely to limit their practicability is a lack of robustness for hyperbolic PDEs, or more broadly for advection-dominated problems. In particular, this is true for the multigrid reduction-in-time (MGRIT) algorithm [15], which is the algorithm that we focus on in this work.

MGRIT enables time-parallelism by applying local time-stepping in parallel across

*Submitted to the editors March 24, 2022.

Funding: This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 (LLNL-JRNL-833088). This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Applied Mathematics program, and by NSERC of Canada.

[†]Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada (hans.desterck@uwaterloo.ca).

[‡]Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, Livermore, California, USA (falgout2@llnl.gov).

[§]Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada (okrzysik@uwaterloo.ca, <https://orcid.org/0000-0001-7880-6512>). The work of this author was partially supported by an Australian Government Research Training Program Scholarship.

a sequence of temporally coarsened grids combined with a global coarse-grid correction on a coarsest grid. The well-known time-parallel Parareal algorithm [28] can be interpreted as a special case of MGRIT that uses only two levels and a specific choice for the relaxation scheme [21], and thus our work also applies to Parareal. MGRIT is iterative, and its convergence properties hinge on how well the coarse-grid problem approximates the fine-grid problem. The default technique for developing the coarse-grid problem is to *rediscretize* the fine-grid PDE on the coarsened mesh, and this tends to work excellently for diffusion-dominated problems [28, 1, 15, 17, 16]. Conversely, it is well-documented that convergence tends to be substantially worse for advection-dominated PDEs when using rediscretization, or closely related techniques [7, 8, 12, 22, 18, 23, 25, 24, 31, 34, 35, 41, 11, 10, 27].

Unsurprisingly, it has also been observed that MGRIT convergence may deteriorate substantially when transitioning from the diffusion-dominated regime of an advection-diffusion problem to the advection-dominated regime [41, 24, 35, 34]. In fact, this observation served as the motivation for an idea that was explored in [36], with some preliminary numerical experiments showing that MGRIT convergence of certain advection discretizations can be improved substantially by adding judiciously chosen amounts of numerical dissipation on the fine grid.

While we focus on multigrid-in-time methods in this paper, we note that some other parallel-in-time methods have shown potential for solving advection-dominated or wave-related problems. Some examples include using reduction-based algebraic multigrid on the space-time system [37], block-circulant preconditioning techniques applied to the space-time system [29], and direct solution techniques based on diagonalizing in the time direction [20].

In this paper, we develop a novel coarse-grid operator leading to the fast MGRIT solution of linear advection problems, building on our earlier work [10] of optimizing coarse-grid operators for constant-wave-speed advection problems. In that work, we showed that it is possible to achieve fast MGRIT convergence on advection problems when using a carefully chosen coarse-grid operator. The coarse-grid operator proposed here is based on a semi-Lagrangian discretization, and is applied to certain classes of semi-Lagrangian discretizations on the fine grid; unlike the coarse-grid operators in [10], the coarse-grid operator is practically computable, and not limited only to constant-wave-speed problems. Employing a semi-Lagrangian coarse-grid operator was tested in [35]; however, solver convergence was not robust with respect to the strength of advection in the advection-diffusion test problem considered. In fact, we will show even for the case of constant-wave-speed advection that simply employing a coarse-grid semi-Lagrangian operator does not result in effective convergence.

Our semi-Lagrangian coarse-grid operator is designed to provide a more accurate approximation to the ideal coarse-grid operator than basic rediscretization, using a truncation error approach to more faithfully match the coarse operator to the ideal coarse operator. Our motivation stems in part from the realization that MGRIT convergence issues for advection-dominated problems are in some sense analogous to the well-known issues that plague the multigrid solution of steady state advection-dominated problems [3, 4, 46, 42, 43, 2]. In particular, the idea of increasing the order of accuracy of the coarse-grid operator for steady state advection relative to the fine-grid operator, as first proposed in [46] and pointed out to us by the author, is a key insight and inspiration for our work. Developing such a coarse-grid operator in the MGRIT context involves additional stability challenges due to the coarsening being in only one coordinate direction rather than in all directions uniformly. The link between slow multigrid convergence in the case of steady state advection in multi-

dimensional space, as analyzed in [46], and slow multigrid convergence in the case of space-time hyperbolic problems with MGRIT, is investigated in detail in [27, Chap. 3].

This manuscript focuses on the MGRIT solution of semi-Lagrangian discretizations of advection problems; however, we stress that the principles developed here can be extended to the MGRIT solution of classical method-of-lines discretizations of hyperbolic problems, which is a subject of our on-going research. Finally, we remark that this manuscript is based on the PhD thesis [27, Chap. 4], and as such, we refer the interested reader there for further details on this work.

The remainder of this manuscript is organized as follows. [Section 2](#) describes the semi-Lagrangian discretizations we consider, followed by an algorithmic description of MGRIT and motivating examples. [Section 3](#) presents the proposed coarse-grid operator. [Section 4](#) generalizes the coarse-grid operator from one to two spatial dimensions. [Section 5](#) presents parallel results. Concluding remarks are given in [Section 6](#).

2. Preliminaries.

2.1. Semi-Lagrangian discretization of the one-dimensional advection equation. We consider semi-Lagrangian discretizations of one-dimensional, variable-wave-speed advection problems of the form

$$(2.1) \quad \frac{\partial u}{\partial t} + \alpha(x, t) \frac{\partial u}{\partial x} = 0, \quad (x, t) \in \Omega \times (0, T], \quad u(x, 0) = u_0(x),$$

for spatial domain $\Omega \subset \mathbb{R}$, and solution u subject to periodic boundary conditions on $\partial\Omega$. Specifically, our numerical tests for this one-dimensional problem will use the initial condition $u_0(x) = \sin^4(\pi x)$, and the spatial domain $\Omega = (-1, 1)$. The MGRIT coarse-grid operators we propose for the multigrid solution of these problems rely on a detailed understanding of semi-Lagrangian methods, which we now describe. We consider semi-Lagrangian discretizations of (2.1) that are based on finite differences, assume a sufficient degree of smoothness of the solution, and assume a readily computable wave-speed $\alpha(x, t)$ for any x and t is available. The reader is directed to [14] and [13, Sec. 7] for more detailed descriptions of semi-Lagrangian methods.

Consider discretizing the spatial domain Ω with a set of n_x equidistant nodes $\mathbf{x} = (x_1, \dots, x_{n_x})^\top$, $x_{i+1} = x_i + h$. Furthermore, discretize the time interval $t \in [0, T]$ with an equidistant mesh of $n_t + 1$ nodes $0 = t_0 < \dots < t_{n_t} = T$, $t_{n+1} = t_n + \delta t$. Given the vector $\mathbf{u}_n \approx u(\mathbf{x}, t_n)$, which represents the approximate solution of (2.1) at time t_n at spatial mesh points \mathbf{x} , the semi-Lagrangian method advances this to a new approximation $\mathbf{u}_{n+1} \approx u(\mathbf{x}, t_{n+1})$ at t_{n+1} as we now describe.

The Lagrangian formulation of the PDE (2.1) is

$$(2.2) \quad \frac{d}{dt} \xi(t) = \alpha(\xi(t), t), \quad \frac{d}{dt} u(\xi(t), t) = 0,$$

in which the curves $(x, t) = (\xi(t), t)$ are the *characteristics* of the PDE. On some characteristic $\xi_i(t)$, the evolution equation—the latter equation in (2.2)—states that the PDE solution is constant. Since the solution at the mesh point $(x, t) = (x_i, t_{n+1})$ is desired, one forces the characteristic $\xi_i(t)$ to pass through this point, such that the solution at this point is simply the solution at the foot of the characteristic at $t = t_n$ (see [Figure 1](#)). To this end, define the local characteristic $\xi_i^{(t_n, \delta t)}(t)$ to be that which passes through the *arrival point* $(x, t) = (x_i, t_{n+1})$. Then, the associated *departure*

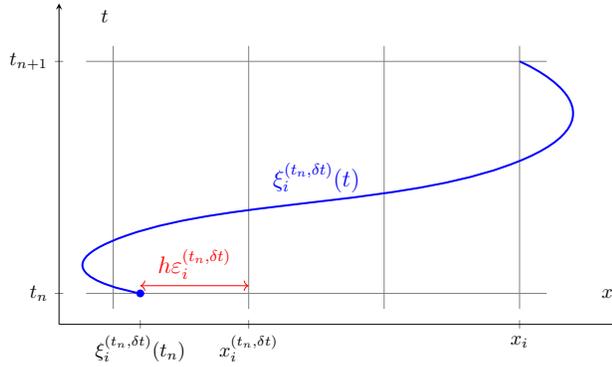


FIG. 1. A local characteristic $\xi_i^{(t_n, \delta t)}(t)$ for $t \in [t_n, t_{n+1}]$ of the variable-wave-speed problem (2.1). By definition, the characteristic passes through the arrival point $(x, t) = (x_i, t_{n+1})$. The departure point (or foot) of the characteristic is its location at time $t = t_n$. The departure point is decomposed into the sum of its east-neighboring mesh point and its distance from this point (see (2.4)).

point $(x, t) = (\xi_i^{(t_n, \delta t)}(t_n), t_n)$ is given by the solution of the final-value problem

$$(2.3) \quad \frac{d}{dt} \xi_i^{(t_n, \delta t)}(t) = \alpha(\xi_i^{(t_n, \delta t)}(t), t), \quad t \in [t_n, t_{n+1}], \quad \xi_i^{(t_n, \delta t)}(t_{n+1}) = x_i,$$

at $t = t_n$. For general wave-speed functions α , the initial-time solution of (2.3) cannot be found exactly, and thus needs to be approximated in some way. In this paper, we approximate this departure point on the fine grid by integrating backwards using a single explicit Runge-Kutta (ERK) step of size δt .¹ Suppose that the ERK method has a global accuracy of order r , such that each departure point is located with an accuracy of $\mathcal{O}(\delta t^{r+1})$ (e.g., forward Euler has $r = 1$). In particular, we will consider ERK schemes with $r = 1, 3$, and 5 , with the Butcher tableaux for the specific schemes available from [27, App. A.1].

Upon (approximately) locating the departure point $\xi_i^{(t_n, \delta t)}(t)$, it will not, in general, coincide with a mesh point, yet the PDE approximation \mathbf{u}_n is only available at mesh points. To resolve this, an interpolating polynomial of at most degree $p \in \mathbb{N}$ is fit through the entries of \mathbf{u}_n at the $p + 1$ mesh nodes nearest to the departure point. Specifically, the departure point is decomposed as

$$(2.4) \quad \xi_i^{(t_n, \delta t)}(t_n) \equiv x_i^{(t_n, \delta t)} - h\varepsilon_i^{(t_n, \delta t)}, \quad \varepsilon_i^{(t_n, \delta t)} \in [0, 1),$$

in which $x_i^{(t_n, \delta t)}$ is the mesh node immediately east of the departure point $\xi_i^{(t_n, \delta t)}(t_n)$, and $\varepsilon_i^{(t_n, \delta t)}$ is its (mesh-normalized) distance from this point; see Figure 1 for a schematic example. The $p + 1$ interpolation nodes are thus $\{x_i^{(t_n, \delta t)} + hj\}_{j=-\ell(p)}^{r(p)}$, with $\ell(p)$ and $r(p)$ —the west and east extents of the stencil, respectively—chosen so that the set of interpolation nodes represent the $p + 1$ nearest neighbors of the departure point.² When p is odd, this results in $\ell(p) = \frac{p+1}{2}$ and $r(p) = \frac{p-1}{2} = \ell(p) - 1$.

¹On coarser levels in our multigrid-in-time hierarchy, we will explore several alternative possibilities rather than simply redeploying the ERK method with a coarse time step, which is not sufficiently accurate.

²For interpolating polynomials of degree $p \geq 1$, both the west and east neighboring mesh points

When p is even, the stencil has a one-point bias, such that $\ell(p)$ and $r(p)$ depend on whether $\varepsilon_i^{(t_n, \delta t)}$ is larger than one half (we ignore this dependence in our notation).

Locating departure points for all arrival points $(x, t) = (\mathbf{x}, t_{n+1})$, then carrying out this piecewise polynomial interpolation constitutes a single time-step of the semi-Lagrangian discretization. We denote the linear time-stepping operator corresponding to this semi-Lagrangian method as $\Phi^{(t_n, \delta t)} = \mathcal{S}_{p,r}^{(t_n, \delta t)} \in \mathbb{R}^{n_x \times n_x}$.

By tracking characteristics (with a sufficient level of accuracy), this discretization ensures the physical domain of dependence lies within the numerical domain of dependence. Generally speaking, this is why semi-Lagrangian discretizations are typically free of a CFL constraint. However, ensuring that characteristics are tracked with sufficient accuracy can in certain circumstances lead to the imposition of a CFL-like constraint, although it is typically looser than that imposed by Eulerian schemes [26, 38]. For sufficiently smooth solutions of (2.1), $\mathcal{S}_{p,r}^{(t_n, \delta t)}$ has a global convergence rate of the form $\mathcal{O}(\delta t^r + \frac{h^{p+1}}{\delta t})$, with the first term arising from approximately locating departure points, and the second from the polynomial interpolation at them [14, Sec. 6.1.2]. Thus, while stability can be maintained with large time steps, they likely lead to a reduction in accuracy of the method. Nonetheless, the time-step size can be chosen based on accuracy requirements rather than stability requirements, unlike CFL-constrained Eulerian schemes for which stability is often the key factor in determining the time-step size. For example, this is why semi-Lagrangian methods see frequent use in numerical weather prediction [30, 40, 44]. In any event, we attempt to balance temporal and spatial errors by using $\delta t \approx h$ and $r = p$, which is common in the literature when developing semi-Lagrangian discretizations [33, 26, 6, 5].

2.2. Algorithmic description of MGRIT. We now provide a brief algorithmic description of MGRIT as it applies to linear problems. Consider the following fully discrete, one-step problem

$$(2.5) \quad \mathbf{u}_{n+1} = \Phi^{(t_n, \delta t)} \mathbf{u}_n + \mathbf{g}_{n+1}, \quad n = 0, 1, \dots, n_t - 1,$$

in which \mathbf{u}_n denotes the approximate spatial solution of some time-dependent PDE at time $t = t_n$, and the initial condition \mathbf{u}_0 is given. The vector \mathbf{g}_n contains solution-independent information, and $\Phi^{(t_n, \delta t)}$ is the (linear) time-stepping operator. Of course, (2.5) may be solved straightforwardly with sequential time-stepping—solving for \mathbf{u}_1 given \mathbf{u}_0 , then solving for \mathbf{u}_2 given \mathbf{u}_1 , and so on, but this process is inherently sequential. In contrast, MGRIT solves (2.5) for all unknowns $\mathbf{u}_n, n = 1, \dots, n_t$, at once in parallel using multigrid reduction techniques, as we now detail.

Suppose that the underlying time grid in (2.5) is equispaced with $t_{n+1} = t_n + \delta t$. Then, let a coarsening factor $m \in \mathbb{N}$ induce a coarse grid defined by taking every m th time point from the fine grid. Define F-points as those appearing exclusively on the fine grid, and all other points as C-points. Given an approximate solution of (2.5), define a C-relaxation of (2.5) as updating C-point values to have zero residuals—this is achieved by time-stepping to each C-point from the F-point immediately before it. Further, define an F-relaxation of (2.5) as updating F-point values to have zero residuals—this is achieved by sequentially time-stepping from each C-point across the $m - 1$ F-points immediately after it.

appear in the interpolation stencil of the departure point. Therefore, the choice made in (2.4) to write the departure point in terms of its east neighbor $x_i^{(t_n, \delta t)}$ is arbitrary in the sense that it could also have been written in terms of its west neighbor.

Given an approximate solution of (2.5), a two-level MGRIT iteration proceeds by pre-relaxation, coarse-grid correction, and post-relaxation. The typical pre-relaxation scheme is FCF: An F-relaxation followed by a C-relaxation followed by a further F-relaxation. However, a single F-relaxation can also be used, as is often the case in the Parareal literature [21]. In this work, we exclusively use FCF-relaxation, since we find it often results in more robust convergence for the advection problems we consider—in some cases, it can even mean the difference between a quickly converging solver and a divergent solver. Post-relaxation is simply an F-relaxation. Note that the pre- and post-relaxation sweeps are highly parallelizable. The coarse-grid problem is derived from the C-point Schur complement of the residual equation of (2.5). More specifically, one computes an approximate error at C-points by solving the following problem containing fewer time points:

$$(2.6) \quad \mathbf{e}_{m(n+1)} = \Phi^{(t_{mn}, m\delta t)} \mathbf{e}_{mn} + \mathbf{r}_{m(n+1)}, \quad n = 0, 1, \dots, (n_t - 1)/m,$$

in which $\mathbf{r}_{m(n+1)}$ is the algebraic residual of (2.5) at the C-point $t = t_{m(n+1)}$. Here $\Phi^{(t_{mn}, m\delta t)}$ is the *coarse-grid time-stepping operator*, and it should approximate the *ideal* coarse-grid time-stepping operator defined by stepping across the coarse time interval $t \in [t_{mn}, t_{mn} + m\delta t]$ using the fine-grid operators: $\Phi^{(t_{mn}, m\delta t)} \approx \prod_{k=0}^{m-1} \Phi^{(t_{mn+k}, \delta t)} =: \Phi_{\text{ideal}}^{(t_{mn}, m\delta t)}$. Upon solving (2.6), the approximate coarse-grid error is interpolated to the fine grid via injection—i.e., added to existing C-point values. Equation (2.6) may be solved by sequential time-stepping, resulting in a two-level method. Alternatively, since (2.6) has the same structure as (2.5), its solution may be approximated in parallel by applying the MGRIT algorithm recursively, resulting in a multilevel method.

In the event that $\Phi^{(t_{mn}, m\delta t)} = \Phi_{\text{ideal}}^{(t_{mn}, m\delta t)}$, (2.6) is exactly the C-point Schur complement of the residual equation associated with (2.5) and MGRIT converges in a single iteration. However, stepping across the coarse time interval $t \in [t_{mn}, t_{mn} + m\delta t]$ with $\Phi_{\text{ideal}}^{(t_{mn}, m\delta t)}$ is just as expensive as stepping across this interval on the fine grid, and thus no parallel speed-up can be achieved. Therefore, stepping with $\Phi^{(t_{mn}, m\delta t)}$ should be less expensive than the ideal coarse-grid operator. Crucially, however, fast convergence of the method generally requires that the coarse-grid operator accurately approximates (in a certain sense) its ideal counterpart [12, 39, 10].

Our numerical tests will use the open-source MGRIT implementation provided by XBraid [45]. In our tests, the initial MGRIT iterate will be taken as a vector with entries uniformly random between zero and one. The metric that we use to report convergence is the number of MGRIT iterations required to reduce the space-time residual by at least 10 orders of magnitude in the two-norm from its initial value.

2.3. Motivating examples. As mentioned previously, a discretized coarse-grid operator often yields fast convergence for diffusion-dominated problems, but tends to be a poor choice for advection-dominated problems [27, 10]. We now demonstrate that a discretized coarse-grid semi-Lagrangian operator is indeed also a poor choice for our model problem. To do so, we consider PDE (2.1) with constant wave-speed $\alpha = 1$, and examine the convergence factor of two-level MGRIT. Specifically, for fixed coarsening factor m , we consider the function

$$(2.7) \quad \rho(c) = \max_{\omega \in [-\pi, \pi)} |\lambda(\omega; c)|^m \frac{|\lambda^m(\omega; c) - \mu(\omega; c)|}{1 - |\mu(\omega; c)|},$$

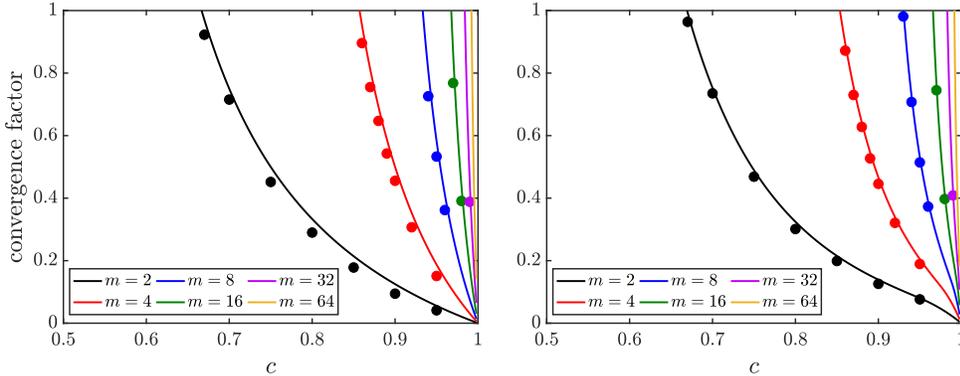


FIG. 2. Convergence factor for two-level MGRIT to solve semi-Lagrangian discretizations of the constant-wave-speed advection problem (2.1) using a rediscretized semi-Lagrangian coarse-grid operator. Left: semi-Lagrangian discretization order $p = 1$; right: $p = 3$. For a fixed coarsening factor m , solid lines show the Fourier analysis estimate of the asymptotic convergence factor (2.7). Solid markers show convergence factors measured on the final iteration of numerical experiments.

in which $c = \delta t/h$ is the (fine-grid) CFL number. For fixed c , the quantities $\lambda(\omega; c)$ and $\mu(\omega; c)$ in (2.7) are the Fourier symbols of the fine- and coarse-grid time stepping operators, respectively, as functions of spatial Fourier frequency ω . For fixed c , (2.7) is the Fourier analysis estimate of the asymptotic convergence factor of two-level MGRIT as $n_t \rightarrow \infty$ [27, Chap. 3]. Provided the number of iterations is not so large that the initial condition has been sequentially propagated across much of the time-domain via the fine-grid relaxation scheme—the practically relevant case for MGRIT—, (2.7) provides a useful estimate of the MGRIT convergence factor for finite n_t .

For semi-Lagrangian orders $p = 1, 3$, (2.7) is plotted in Figure 2 for $c \in [0.5, 1]$.³ For a given c , to numerically evaluate the maximum in (2.7) over ω we use 512 equispaced points in $[-\pi, \pi)$. In Figure 2 for each m , there is a significant interval of CFL numbers for which $\rho > 1$, indicating that the residual norm grows from one iteration to the next (that is, at least in iterations for which (2.7) is a valid estimate). Moreover, the size of interval for which $\rho > 1$ appears to grow with m . We conclude that simply rediscretizing the fine-grid semi-Lagrangian scheme on the coarse grid does not result in a robust MGRIT solver for our model problem. We do not provide analogous plots for even polynomial degrees $p = 2, 4$ because the convergence factor (2.7) is larger than one for almost all c in these cases.

To demonstrate that (2.7) is an accurate estimate of the true convergence factor, overlaid on the plots in Figure 2 are experimentally measured convergence factors taken from the final MGRIT iteration in numerical experiments. The tests use a space-time domain with $n_x \times n_t = 2^7 \times 2^{15}$ points. Note also that analogous numerical experiments for variable-wave-speed problems indicate that rediscretization is a poor choice for those problems too.

3. Coarse-grid operator in one spatial dimension. From Figure 2, it is clear that, in general, rediscretizing the semi-Lagrangian operator on the coarse grid does not yield robust MGRIT convergence. Despite this, the idea of using a semi-

³Specifically for the constant-wave-speed problem, by symmetry of the fine- and coarse-grid operators used here when p is odd, the corresponding plots for CFL numbers $c \in [0, 0.5]$ are a reflection of those for $c \in [0.5, 1]$ in Figure 2 about $c = 0.5$. Similarly, the corresponding plots for $c \in [k, k + 1]$ with $k \in \mathbb{N}$ are identical to those for $c \in [0, 1]$.

Lagrangian discretization on time-coarsened grids remains appealing because the discretization is stable for all time-step sizes. Furthermore, in [10], we argued both for stability, and in terms of approximating the ideal coarse-grid operator, that a coarse-grid operator for an advection-dominated problem should be semi-Lagrangian-like in nature, in the sense that its numerical domain of dependence should track the characteristic curves. For these reasons, we seek a coarse-grid operator that is based on a semi-Lagrangian discretization, but that provides a better approximation to the ideal coarse-grid operator than the semi-Lagrangian discretization does. There are many possible metrics one could use to characterize the difference between a coarse-grid operator and its ideal counterpart. Here, we use the concept of local truncation error, which is defined as the amount by which the exact PDE solution fails to satisfy the discrete scheme after one time step. Accordingly, we begin in the following section with estimates for the truncation error of the fine-grid semi-Lagrangian discretization, and the associated ideal coarse-grid operator.

Throughout the rest of the paper, the matrix $\mathcal{D}_{p+1} \in \mathbb{R}^{n_x \times n_x}$ is defined such that $h^{-(p+1)}\mathcal{D}_{p+1}$ represents a finite-difference rule for approximating the $p+1$ st derivative of periodic grid functions. Let $\mathbf{v} = (v(x_1), \dots, v(x_{n_x}))^\top \in \mathbb{R}^{n_x}$ denote a vector of a periodic function $v(x)$ evaluated on the spatial mesh. Then, if the finite-difference rule is of order $s \in \mathbb{N}$ and v is at least $p+1+s$ times continuously differentiable

$$(3.1) \quad \left(\frac{\mathcal{D}_{p+1}}{h^{p+1}} \mathbf{v} \right)_i = \frac{d^{p+1}v}{dx^{p+1}} \Big|_{x_i} + \mathcal{O}(h^s), \quad i \in \{1, \dots, n_x\}.$$

The order s is not particularly important for our purposes so we do not specify it, but note that numerical tests in later sections use finite-difference approximations with $s = 2$. Since the mesh points are equispaced, the matrix \mathcal{D}_{p+1} is circulant. Finally, note that while the entries of \mathcal{D}_{p+1} are independent of h , its action is not: $\mathcal{D}_{p+1}\mathbf{v} = \mathcal{O}(h^{p+1})$ if \mathbf{v} is independent of h .

3.1. Truncation error estimates for exact departure points. For simplicity, consider an idealized semi-Lagrangian discretization $\mathcal{S}_{p,r}^{(t_n, \delta t)}$ that locates departure points exactly, which we will denote by writing $r = \infty$. Furthermore, the following error estimates (Lemma 3.1, Corollary 3.2, and Lemma 3.3) assume that the wave-speed function may depend on time but is independent of space. Discussion on spatially varying wave-speed functions is given after Lemma 3.3. Note that [14, p. 170] considers a simpler but related type of error estimate to Lemma 3.1 when α is constant.

LEMMA 3.1 (Semi-Lagrangian truncation error for $r = \infty$). *Let \mathcal{D}_{p+1} be as in (3.1). Suppose that the solution $u(x, t)$ of (2.1) is at least $p+1+s$ times continuously differentiable with respect to x , and that the wave-speed in (2.1) is independent of space, $\alpha(x, t) \equiv \alpha(t)$. Define $\mathbf{u}(t) \in \mathbb{R}^{n_x}$ as the vector composed of the PDE solution sampled in space at the mesh points \mathbf{x} and at time t , and let $\mathcal{S}_{p,\infty}^{(t_n, \delta t)}$ be the semi-Lagrangian discretization of (2.1) that locates departure points exactly at time t_n . Then, the local truncation error of this discretization can be expressed as*

$$(3.2) \quad \mathbf{u}(t_{n+1}) - \mathcal{S}_{p,\infty}^{(t_n, \delta t)} \mathbf{u}(t_n) = (-h)^{p+1} \text{diag} \left(f_{p+1}(\boldsymbol{\varepsilon}^{(t_n, \delta t)}) \right) \frac{\mathcal{D}_{p+1}}{h^{p+1}} \mathbf{u}(t_{n+1}) + \mathcal{O}(h^{p+2}),$$

in which $\boldsymbol{\varepsilon}^{(t_n, \delta t)} := (\varepsilon_1^{(t_n, \delta t)}, \dots, \varepsilon_{n_x}^{(t_n, \delta t)})^\top \in \mathbb{R}^{n_x}$, and $\varepsilon_i^{(t_n, \delta t)}$ is defined in (2.4). The

function f_{p+1} in (3.2) is the following degree $p+1$ polynomial

$$(3.3) \quad f_{p+1}(z) := \frac{1}{(p+1)!} \prod_{q=-\ell(p)}^{r(p)} (q+z).$$

Note that since $\mathbf{u}(t_{n+1})$ is independent of h , the vector $\frac{\mathcal{D}_{p+1}}{h^{p+1}}\mathbf{u}(t_{n+1})$ in (3.2) is independent of h to leading order (see (3.1)).

Proof. Since $\mathcal{S}_{p,\infty}^{(t_n,\delta t)}$ exactly locates departure points of (2.1), the only truncation error resulting from applying the i th row of $\mathcal{S}_{p,\infty}^{(t_n,\delta t)}$ to $u(\mathbf{x}, t_n)$ is the error from the polynomial interpolation at the departure point $\xi_i^{(t_n,\delta t)}(t_n) = x_i^{(t_n,\delta t)} - h\varepsilon_i^{(t_n,\delta t)}$. Since u is at least $p+1$ times continuously differentiable with respect to x , the standard error estimate from polynomial interpolation theory can be applied (see, e.g., [9, Thm. 3.1.1]). Since the interpolation nodes are separated by distance h , applying this estimate at the i th departure point at time t_n yields

$$(3.4) \quad \begin{aligned} & u(\xi_i^{(t_n,\delta t)}(t_n), t_n) - \left(\mathcal{S}_{p,\infty}^{(t_n,\delta t)} \mathbf{u}(t_n) \right)_i \\ &= \frac{1}{(p+1)!} \prod_{q=-\ell(p)}^{r(p)} \left[(x_i^{(t_n,\delta t)} - h\varepsilon_i^{(t_n,\delta t)}) - (x_i^{(t_n,\delta t)} + hq) \right] \frac{\partial^{p+1}u}{\partial x^{p+1}} \Big|_{(\zeta_i^{(t_n,\delta t)}, t_n)} \end{aligned}$$

$$(3.5) \quad = (-h)^{p+1} f_{p+1}(\varepsilon_i^{(t_n,\delta t)}) \frac{\partial^{p+1}u}{\partial x^{p+1}} \Big|_{(\zeta_i^{(t_n,\delta t)}, t_n)},$$

for some unknown point $\zeta_i^{(t_n,\delta t)} \in (x_i^{(t_n,\delta t)} - h\ell(p), x_i^{(t_n,\delta t)} + hr(p))$. Since $\zeta_i^{(t_n,\delta t)}$ and $\xi_i^{(t_n,\delta t)}(t_n)$ are a distance of $\mathcal{O}(h)$ apart, write $\zeta_i^{(t_n,\delta t)} = \xi_i^{(t_n,\delta t)}(t_n) + h\widehat{\zeta}_i^{(t_n,\delta t)}$ for some other unknown quantity $\widehat{\zeta}_i^{(t_n,\delta t)}$. Then, by Taylor expansion,

$$(3.6) \quad \frac{\partial^{p+1}u}{\partial x^{p+1}} \Big|_{(\zeta_i^{(t_n,\delta t)}, t_n)} = \frac{\partial^{p+1}u}{\partial x^{p+1}} \Big|_{(\xi_i^{(t_n,\delta t)}(t_n) + h\widehat{\zeta}_i^{(t_n,\delta t)}, t_n)} = \frac{\partial^{p+1}u}{\partial x^{p+1}} \Big|_{(\xi_i^{(t_n,\delta t)}(t_n), t_n)} + \mathcal{O}(h).$$

Next, observe that differentiating the advection problem (2.1) $p+1$ times with respect to x yields when $\alpha(x, t) \equiv \alpha(t)$

$$(3.7) \quad \frac{\partial}{\partial t} \frac{\partial^{p+1}u}{\partial x^{p+1}} + \alpha(t) \frac{\partial}{\partial x} \frac{\partial^{p+1}u}{\partial x^{p+1}} = 0.$$

Thus, $\frac{\partial^{p+1}u}{\partial x^{p+1}}$ is constant along characteristics of the advection problem (2.1), from which it follows that $\frac{\partial^{p+1}u}{\partial x^{p+1}} \Big|_{(\xi_i^{(t_n,\delta t)}(t_n), t_n)} = \frac{\partial^{p+1}u}{\partial x^{p+1}} \Big|_{(x_i, t_{n+1})}$. Applying this in (3.6) and substituting the result along with $u(\xi_i^{(t_n,\delta t)}(t_n), t_n) = u(x_i, t_{n+1})$ into (3.5) gives

$$(3.8) \quad u(x_i, t_{n+1}) - \left(\mathcal{S}_{p,\infty}^{(t_n,\delta t)} \mathbf{u}(t_n) \right)_i = (-h)^{p+1} f_{p+1}(\varepsilon_i^{(t_n,\delta t)}) \frac{\partial^{p+1}u}{\partial x^{p+1}} \Big|_{(x_i, t_{n+1})} + \mathcal{O}(h^{p+2}).$$

Finally, substituting into this equation $\frac{\partial^{p+1}u}{\partial x^{p+1}} \Big|_{(x_i, t_{n+1})} = \left(\frac{\mathcal{D}_{p+1}}{h^{p+1}} \mathbf{u}(t_{n+1}) \right)_i + \mathcal{O}(h^s)$ (see (3.1)) gives the i th row of the claimed result (3.2). \square

COROLLARY 3.2 (Ideal coarse-grid semi-Lagrangian truncation error for $r = \infty$).

Suppose the assumptions of Lemma 3.1 hold. Then, the ideal coarse-grid operator defined by time-stepping across $t \in [t_n, t_n + m\delta t]$ with the m fine-grid operators $\mathcal{S}_{p,\infty}^{(t_n+k\delta t, \delta t)}$, $k = 0, 1, \dots, m-1$, has a local truncation error given by

$$(3.9) \quad \begin{aligned} \mathbf{u}(t_{n+m}) - \left[\prod_{k=0}^{m-1} \mathcal{S}_{p,\infty}^{(t_n+k\delta t, \delta t)} \right] \mathbf{u}(t_n) \\ = (-h)^{p+1} \text{diag} \left(\sum_{k=0}^{m-1} f_{p+1}(\boldsymbol{\varepsilon}^{(t_n+k\delta t, \delta t)}) \right) \frac{\mathcal{D}_{p+1}}{h^{p+1}} \mathbf{u}(t_{n+m}) + \mathcal{O}(h^{p+2}). \end{aligned}$$

Proof. Applying $\mathcal{S}_{p,\infty}^{(t_{n+1}, \delta t)}$ to both sides of (3.2) gives

$$(3.10) \quad \begin{aligned} \mathcal{S}_{p,\infty}^{(t_{n+1}, \delta t)} \mathbf{u}(t_{n+1}) - \left[\mathcal{S}_{p,\infty}^{(t_{n+1}, \delta t)} \mathcal{S}_{p,\infty}^{(t_n, \delta t)} \right] \mathbf{u}(t_n) \\ = (-h)^{p+1} \mathcal{S}_{p,\infty}^{(t_{n+1}, \delta t)} \text{diag} \left(f_{p+1}(\boldsymbol{\varepsilon}^{(t_n, \delta t)}) \right) \frac{\mathcal{D}_{p+1}}{h^{p+1}} \mathbf{u}(t_{n+1}) + \mathcal{O}(h^{p+2}), \end{aligned}$$

$$(3.11) \quad = (-h)^{p+1} \text{diag} \left(f_{p+1}(\boldsymbol{\varepsilon}^{(t_n, \delta t)}) \right) \frac{\mathcal{D}_{p+1}}{h^{p+1}} \left(\mathcal{S}_{p,\infty}^{(t_{n+1}, \delta t)} \mathbf{u}(t_{n+1}) \right) + \mathcal{O}(h^{p+2}).$$

To arrive at (3.11), we have used the fact that when the wave-speed is independent of space $\mathcal{S}_{p,\infty}^{(t_{n+1}, \delta t)}$ commutes with $\text{diag} \left(f_{p+1}(\boldsymbol{\varepsilon}^{(t_n, \delta t)}) \right)$ and \mathcal{D}_{p+1} . When the wave-speed is independent of space, characteristics are parallel for all time t , and thus departure points of neighboring local characteristics are equally separated by a distance h , from which it follows that $\boldsymbol{\varepsilon}^{(t_n, \delta t)}$ is constant. Therefore $\text{diag} \left(f_{p+1}(\boldsymbol{\varepsilon}^{(t_n, \delta t)}) \right)$ is a constant diagonal matrix and thus commutes with all other matrices. Furthermore, since neighboring departure points are equispaced, the semi-Lagrangian matrix $\mathcal{S}_{p,\infty}^{(t_{n+1}, \delta t)}$ is circulant, and therefore commutes with the circulant matrix \mathcal{D}_{p+1} .

Moving on, from (3.2) one has

$$(3.12) \quad \begin{aligned} \mathcal{S}_{p,\infty}^{(t_{n+1}, \delta t)} \mathbf{u}(t_{n+1}) \\ = \mathbf{u}(t_{n+2}) - (-h)^{p+1} \text{diag} \left(f_{p+1}(\boldsymbol{\varepsilon}^{(t_{n+1}, \delta t)}) \right) \frac{\mathcal{D}_{p+1}}{h^{p+1}} \mathbf{u}(t_{n+2}) + \mathcal{O}(h^{p+2}). \end{aligned}$$

Substituting this into (3.11) and keeping only terms up to size $\mathcal{O}(h^{p+1})$ gives

$$(3.13) \quad \begin{aligned} \mathbf{u}(\mathbf{x}, t_{n+2}) - \left[\mathcal{S}_{p,\infty}^{(t_{n+1}, \delta t)} \mathcal{S}_{p,\infty}^{(t_n, \delta t)} \right] \mathbf{u}(t_n) = \mathcal{O}(h^{p+2}) + \\ (-h)^{p+1} \text{diag} \left(f_{p+1}(\boldsymbol{\varepsilon}^{(t_n, \delta t)}) + f_{p+1}(\boldsymbol{\varepsilon}^{(t_{n+1}, \delta t)}) \right) \frac{\mathcal{D}_{p+1}}{h^{p+1}} \mathbf{u}(t_{n+2}). \end{aligned}$$

Inductively repeating the above process on (3.13) with the remaining $m-2$ fine-grid operators $\mathcal{S}_{p,\infty}^{(t_n+k\delta t, \delta t)}$, $k = 2, \dots, m-1$, one arrives at the result (3.9). \square

Having developed an asymptotic expansion for the ideal coarse-grid operator, we now relate this to the coarse-grid semi-Lagrangian operator.

LEMMA 3.3 (Perturbed coarse-grid semi-Lagrangian operators). *Suppose the assumptions of Lemma 3.1 hold. Let $\mathcal{S}_{p,\infty}^{(t_n, m\delta t)}$ be the coarse-grid semi-Lagrangian discretization of (2.1) that locates departure points exactly. Then, this operator can be*

expressed as a perturbation of the ideal coarse-grid operator $\prod_{k=0}^{m-1} \mathcal{S}_{p,\infty}^{(t_n+k\delta t, \delta t)}$ in the following three ways:

$$(3.14) \quad \begin{aligned} & \left[\prod_{k=0}^{m-1} \mathcal{S}_{p,\infty}^{(t_n+k\delta t, \delta t)} \right] \mathbf{u}(t_n) \\ &= \mathcal{S}_{p,\infty}^{(t_n, m\delta t)} \mathbf{u}(t_n) + \text{diag} \left(\boldsymbol{\varphi}_{p+1}^{(t_n, m\delta t)} \right) \mathcal{D}_{p+1} \mathbf{u}(t_{n+m}) + \mathcal{O}(h^{p+2}), \end{aligned}$$

$$(3.15) \quad = \left[I + \text{diag} \left(\boldsymbol{\varphi}_{p+1}^{(t_n, m\delta t)} \right) \mathcal{D}_{p+1} \right] \mathcal{S}_{p,\infty}^{(t_n, m\delta t)} \mathbf{u}(t_n) + \mathcal{O}(h^{p+2}),$$

$$(3.16) \quad = \left[I - \text{diag} \left(\boldsymbol{\varphi}_{p+1}^{(t_n, m\delta t)} \right) \mathcal{D}_{p+1} \right]^{-1} \mathcal{S}_{p,\infty}^{(t_n, m\delta t)} \mathbf{u}(t_n) + \mathcal{O}(h^{p+2}),$$

with the vector $\boldsymbol{\varphi}_{p+1}^{(t_n, m\delta t)} \in \mathbb{R}^{n_x}$ defined by

$$(3.17) \quad \boldsymbol{\varphi}_{p+1}^{(t_n, m\delta t)} := (-1)^{p+1} \left(f_{p+1}(\boldsymbol{\varepsilon}^{(t_n, m\delta t)}) - \sum_{k=0}^{m-1} f_{p+1}(\boldsymbol{\varepsilon}^{(t_n+k\delta t, \delta t)}) \right).$$

Proof. From the truncation error of the fine-grid semi-Lagrangian discretization $\mathcal{S}_{p,\infty}^{(t_n, \delta t)}$ given in (3.2), it can immediately be seen that the truncation of the coarse-grid semi-Lagrangian discretization $\mathcal{S}_{p,\infty}^{(t_n, m\delta t)}$ is

$$(3.18) \quad \begin{aligned} & \mathbf{u}(t_{n+m}) - \mathcal{S}_{p,\infty}^{(t_n, m\delta t)} \mathbf{u}(\mathbf{x}, t_n) \\ &= (-h)^{p+1} \text{diag} \left(f_{p+1}(\boldsymbol{\varepsilon}^{(t_n, m\delta t)}) \right) \frac{\mathcal{D}_{p+1}}{h^{p+1}} \mathbf{u}(t_{n+m}) + \mathcal{O}(h^{p+2}). \end{aligned}$$

Equation (3.14) follows by subtracting the truncation error (3.18) from that of the ideal coarse-grid operator's in (3.9), and then rearranging the resulting equation for the ideal coarse-grid operator. Equation (3.15) follows by substituting $\mathbf{u}(t_{n+m}) = \mathcal{S}_{p,\infty}^{(t_n, m\delta t)} \mathbf{u}(\mathbf{x}, t_n) + \mathcal{O}(h^{p+1})$, as is given by (3.18), into (3.14). Finally, (3.16) follows from the geometric expansion $[I - \text{diag}(\boldsymbol{\varphi}_{p+1}^{(t_n, m\delta t)}) \mathcal{D}_{p+1}]^{-1} = I + \text{diag}(\boldsymbol{\varphi}_{p+1}^{(t_n, m\delta t)}) \mathcal{D}_{p+1} + (\text{diag}(\boldsymbol{\varphi}_{p+1}^{(t_n, m\delta t)}) \mathcal{D}_{p+1})^2 + \dots$, and, so, $[I - \text{diag}(\boldsymbol{\varphi}_{p+1}^{(t_n, m\delta t)}) \mathcal{D}_{p+1}]^{-1} \mathbf{v} = [I + \text{diag}(\boldsymbol{\varphi}_{p+1}^{(t_n, m\delta t)})] \mathbf{v} + \mathcal{O}(h^{2(p+1)})$ for sufficiently smooth \mathbf{v} . \square

Remark 3.4 (Estimates for spatially varying wave-speed functions). The estimates in Lemma 3.1, Corollary 3.2, and Lemma 3.3 were derived for PDE (2.1) with $\alpha(x, t) \equiv \alpha(t)$. In particular, the proof of Lemma 3.1 does not generalize to the spatially variable case because then $\frac{\partial^{p+1} u}{\partial x^{p+1}}$ is no longer constant along characteristics of (2.1). Furthermore, the commutativity of $\mathcal{S}_{p,\infty}^{(t_n, \delta t)}$ with $\text{diag}(f_{p+1}(\boldsymbol{\varepsilon}^{(t_n, \delta t)}))$ and \mathcal{D}_{p+1} used in the proof of Corollary 3.2 no longer holds. However, in what follows below, we find numerically that the MGRIT method that is derived from the estimates in Lemma 3.3 also works effectively for the case that α depends on x and p is odd. For odd p , we conjecture based on numerical evidence reported in Supplementary Materials Section SM1, that the estimates for the spatially independent wave-speed case hold for the spatially variable case up to terms of size $\mathcal{O}(h^{p+1}\delta t)$, while for even p , we do not believe that all of the estimates extend analogously.

3.2. Coarse-grid operator for exact departure points. The significance of Lemma 3.3 is that it provides several asymptotic relationships between the ideal coarse-grid operator and the coarse-grid semi-Lagrangian operator. Specifically, considering (3.14), when the wave-speed function does not vary in space, the coarse-grid

semi-Lagrangian operator serves as an $\mathcal{O}(h^{p+1})$ approximation to the ideal coarse-grid operator, recalling that $\mathcal{D}_{p+1}\mathbf{u}(t_{n+m}) = \mathcal{O}(h^{p+1})$ (see (3.1)). However, the operators appearing in (3.15) and (3.16) serve as $\mathcal{O}(h^{p+2})$ approximations to the ideal coarse-grid operator when the wave-speed does not vary in space.

Based on (3.15), we propose the following *explicit* coarse-grid operator

$$(3.19) \quad \Phi^{(t_n, m\delta t)} = \mathcal{F}_{p+1}^{(t_n, m\delta t)} \mathcal{S}_{p, \infty}^{(t_n, m\delta t)}, \quad \mathcal{F}_{p+1}^{(t_n, m\delta t)} := I + \text{diag} \left(\varphi_{p+1}^{(t_n, m\delta t)} \right) \mathcal{D}_{p+1}.$$

Based on (3.16) we propose the following *implicit-explicit* coarse-grid operator

$$(3.20) \quad \Phi^{(t_n, m\delta t)} = \mathcal{B}_{p+1}^{(t_n, m\delta t)} \mathcal{S}_{p, \infty}^{(t_n, m\delta t)}, \quad \mathcal{B}_{p+1}^{(t_n, m\delta t)} := \left[I - \text{diag} \left(\varphi_{p+1}^{(t_n, m\delta t)} \right) \mathcal{D}_{p+1} \right]^{-1}.$$

Based on Remark 3.4, we explore using these coarse-grid operators also for problems in which the wave-speed varies in space, even though this introduces further error terms in (3.15) and (3.16) that we conjecture are of size $\mathcal{O}(h^{p+1}\delta t)$ (at least when p is odd). The $\mathcal{F}_{p+1}^{(t_n, m\delta t)}$ and $\mathcal{B}_{p+1}^{(t_n, m\delta t)}$ notation is used to represent forward and backward Euler steps, respectively. Herein, we will typically refer to (3.19) and (3.20) as the ‘forward Euler’ and ‘backward Euler’ coarse-grid operators, respectively. In our tests, \mathcal{D}_{p+1} will be taken as a periodic, finite-difference approximation with second-order accuracy.

The motivation for this nomenclature is that the coarse-grid operators can be interpreted as particular coarse-grid discretizations of a certain PDE. For simplicity, suppose that the wave-speed is constant $\alpha(x, t) \equiv \alpha$, then the vector $\varphi_{p+1}^{(t_n, m\delta t)} \in \mathbb{R}^{n_x}$ given by (3.17) will be constant and independent of t_n , so, let us denote its entries by $\varphi_{p+1}^{(m\delta t)} \in \mathbb{R}$. In this case, the PDE that (3.19) and (3.20) discretize on the coarse grid is

$$(3.21) \quad \frac{\partial u}{\partial t} + \alpha \frac{\partial u}{\partial x} = \varphi_{p+1}^{(m\delta t)} \frac{h^{p+1}}{m\delta t} \frac{\partial u^{p+1}}{\partial x^{p+1}}.$$

More specifically, they use a mixed discretization of (3.21), in which the coarse-grid semi-Lagrangian method $\mathcal{S}_{p, \infty}^{(m\delta t)}$ deals with the advection term, and then the method of lines is applied to solve the rest of the equation. In doing so, the right-hand side of (3.21) is discretized in space using the matrix \mathcal{D}_{p+1} , and the time derivative on the left-hand side is discretized using forward and backward Euler steps in (3.19) and (3.20), respectively. See [27, Sec. 4.2.3] for further details.

An insightful numerical test case for the proposed coarse-grid operators (3.19) and (3.20) is when α is constant, since then departure points can be located exactly—an assumption made in deriving the truncation error estimates. We now provide some general commentary on the results of our numerical tests for constant α to help motivate the direction of the remainder of this paper.

Case 1: Odd p , forward Euler operator (3.19). For sufficiently small m (e.g., $m = 2$ or $m = 4$), we often obtain a quickly converging MGRIT solver, while for larger m the solver often diverges. In fact, we are able to rigorously prove for the case of constant α that the operator $\mathcal{F}_{p+1}^{(t_n, m\delta t)}$ in (3.19) is unstable, in the sense that $\|\mathcal{F}_{p+1}^{(t_n, m\delta t)}\|_2 > 1$, for sufficiently large m (details on this can be found in [27, Sec. 4.2.4]). Note that stability of the coarse-grid operator is a necessary but not sufficient condition for MGRIT convergence. It is likely that the instability of this operator is correlated with the poor performance we observe in our numerical tests for larger

values of m . Since this instability arises even for moderate values of m , we do not believe this operator can be useful in practice.

Case 2: Odd p , backward Euler operator (3.20). Generally speaking, we find this coarse-grid operator yields robust and fast MGRIT convergence. We are able to rigorously prove for the case of constant α that the operator $\mathcal{B}_{p+1}^{(t_n, m\delta t)}$ in (3.20) is unconditionally stable, in the sense that $\|\mathcal{B}_{p+1}^{(t_n, m\delta t)}\|_2 \leq 1$ for all problem parameters (details can be found in [27, Sec. 4.2.4]).

When p is odd, the stability properties of $\mathcal{F}_{p+1}^{(t_n, m\delta t)}$ and $\mathcal{B}_{p+1}^{(t_n, m\delta t)}$ can be understood intuitively from the perspective that these operators resemble forward and backward Euler time discretizations, respectively, of a dissipative PDE $\frac{\partial v}{\partial t} = \kappa_{p+1}^{(m\delta t)} \frac{\partial^{p+1} v}{\partial x^{p+1}}$, with $\kappa_{p+1}^{(m\delta t)}$ a constant. When $p = 1$ this dissipative PDE is the heat equation, for which it is well-known that forward Euler has poor stability properties, while backward Euler has excellent stability properties.

Case 3: Even p . Our tests indicate that the coarse-grid operators do not yield robust MGRIT convergence. We do not yet have a full understanding of why this is the case, and addressing this remains ongoing work. We speculate that this failure is at least in part due to MGRIT being able to more easily correct dissipative errors compared with dispersive errors (see the analysis of [34]). In addition, for even p , the nodes in the interpolation stencil change as a function of the mesh-normalized distance ε to the east neighbor of the departure point (as explained in Section 2.1), and our proposed coarse-grid operator does not account for this.

Given the above discussion, throughout the remainder of this manuscript we focus on the backward Euler coarse-grid operator (3.20) and consider only semi-Lagrangian discretizations using interpolating polynomials of odd degrees p . As an initial demonstration of the improvement that the modified coarse-grid operator (3.20) offers over the standard semi-Lagrangian coarse-grid operator, in Figure 3 we recreate the convergence factor plots from Figure 2 for the constant-wave-speed advection problem.

Remarkably, the plots in Figure 3 show that the proposed coarse-grid operator leads to a convergent MGRIT solver for all CFL numbers, at least when α is constant.

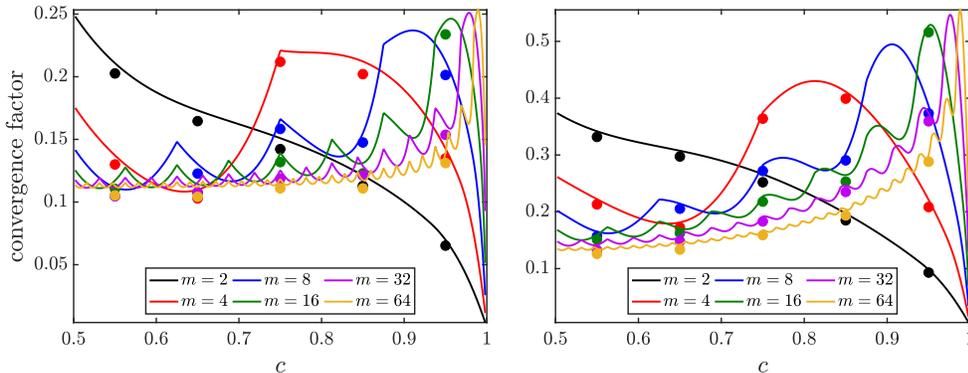


FIG. 3. Convergence factor for two-level MGRIT to solve semi-Lagrangian discretizations of the constant-wave-speed advection problem (2.1) with the proposed backward Euler coarse-grid operator (3.20). Left: semi-Lagrangian discretization order is $p = 1$; right: $p = 3$. For a fixed coarsening factor m , solid lines show the Fourier analysis estimate of the asymptotic convergence factor (2.7) as a function of the CFL number $c = \delta t/h$. Solid markers show experimentally measured convergence factors.

Moreover, convergence is fast for many CFL numbers; however, it does show a somewhat strong dependence on the CFL number and coarsening factor. For a given m , the convergence rate shown in [Figure 3](#) tends to deteriorate going from $p = 1$ to $p = 3$, as is particularly obvious by contrasting the peaks of the $p = 1$ and $p = 3$ curves (note the two different vertical scales). Interestingly, in our prior work [\[10\]](#) on optimizing coarse-grid operators for explicit Eulerian discretizations of constant-wave-speed advection, convergence improved with increasing discretization order. The reversal of this trend here and the non-uniformity of convergence with respect to CFL number perhaps hint that there exist *better* coarse-grid operators than [\(3.20\)](#), but we leave this to future research. In any event, for constant-wave-speed problems, the proposed backward Euler coarse-grid operator results in a robustly converging MGRIT solver for odd polynomial degrees that is fast for the majority of CFL numbers.

A potential concern with the backward Euler operator [\(3.20\)](#) is that applying it requires performing a linear solve, a task which is considerably more computationally expensive than applying an explicit semi-Lagrangian update. Our numerical experiments in the following section will demonstrate, however, that this linear solve may be carried out approximately with a small number of GMRES iterations. In fact, the numerical tests used to generate the data points overlaid in [Figure 3](#) used only 10 GMRES iterations to approximately solve these linear systems.

3.3. Coarse-grid operator for inexact departure points. We now move to the more practical case of developing a coarse-grid operator for when the fine- and coarse-grid semi-Lagrangian methods do not exactly locate departure points. Recall from [Section 2.1](#) that on the fine grid, the semi-Lagrangian method estimates departure points by a single step of an ERK method. The question is now how they should be located on the coarse grid. To answer this, we consider a heuristic strategy, since incorporating inexact departure point locations into the truncation estimates from [Lemma 3.1](#) is not straightforward.

The immediately obvious way to compute coarse-grid departure points is to re-deploy the ERK scheme that was used on the fine grid but with the coarse-grid time-step, i.e., rediscretize the ERK scheme. However, for larger values of m , this can be expected to produce highly inaccurate departure points, at least for variable wave-speeds. Numerical tests for variable-wave-speed problems (not shown here for brevity) confirm that this strategy does not lead to robust MGRIT convergence. A second option (see [\[27, Sec 4.3.2\]](#) for details) is to increase the accuracy of this coarse-grid ERK integration by taking many small steps, such as m steps using the fine-grid time-step δt , for example. The obvious downside of this strategy is that it is expensive, since it uses fine-grid resolution to determine coarse-grid quantities.

Based on the above discussion, for the fine-grid semi-Lagrangian discretization $\mathcal{S}_{p,r}^{(t_n,\delta t)}$ we propose the following backward Euler coarse-grid operator

$$(3.22) \quad \Phi^{(t_n,m\delta t)} = \mathcal{B}_{p+1}^{(t_n,m\delta t)} \mathcal{S}_{p,r_*}^{(t_n,m\delta t)}.$$

Here r_* signifies that the coarse-grid semi-Lagrangian operator $\mathcal{S}_{p,r_*}^{(t_n,m\delta t)}$ should locate departure points with an accuracy that is in some sense comparable to that of the fine-grid operator $\mathcal{S}_{p,r}^{(t_n,\delta t)}$, see [Section 3.3.1](#). The backward Euler matrix $\mathcal{B}_{p+1}^{(t_n,m\delta t)}$ is still defined as it was in [\(3.20\)](#). In [Section 3.3.1](#), a scalable strategy is presented for estimating coarse-grid departure points by reusing the departure point calculations from the fine grid. However, since this discussion is detailed, we first present numerical experiments that use this strategy for the coarse-grid operator [\(3.22\)](#).

		$\alpha = (3.23)$			$\alpha = (3.24)$			$\alpha = (3.25)$		
		m			m			m		
p, r	$n_x \times n_t$	4	8	16	4	8	16	4	8	16
1, 1	$2^8 \times 2^{10}$	14	12	11	12	10	11	11	11	12
	$2^{10} \times 2^{12}$	14	12	11	12	11	11	12	12	12
	$2^{12} \times 2^{14}$	14	12	11	14	11	11	13	13	13
3, 3	$2^8 \times 2^{10}$	22	17	15	15	13	12	13	15	15
	$2^{10} \times 2^{12}$	22	17	15	16	14	13	16	16	16
	$2^{12} \times 2^{14}$	22	17	15	21	15	14	19	19	19
5, 5	$2^8 \times 2^{10}$	30	22	18	18	15	14	15	18	16
	$2^{10} \times 2^{12}$	31	23	20	18	16	15	18	19	20
	$2^{12} \times 2^{14}$	31	23	20	27	18	15	23	24	24

TABLE 1

Number of two-level MGRIT iterations to reach convergence. The wave-speed $\alpha(x, t)$ of the advection problem (2.1) is indicated in the top row of the table. The fine-grid semi-Lagrangian discretization is $\mathcal{S}_{p,r}^{(t_n, \delta t)}$, and the coarse-grid operator is the dissipatively corrected semi-Lagrangian operator (3.22). The coarse-grid semi-Lagrangian operators $\mathcal{S}_{p,r_*}^{(t_n, m\delta t)}$ estimate departure points using the linear interpolation and backtracking strategy from Section 3.3.1.

In the numerical tests, the fine-grid time-step is chosen as $\delta t = 0.85h$, and the following wave-speed functions are considered

$$(3.23) \quad \alpha(x, t) = 1,$$

$$(3.24) \quad \alpha(x, t) = \cos(2\pi t),$$

$$(3.25) \quad \alpha(x, t) = \cos(2\pi t) \cos(2\pi x).$$

Plots of the functions (3.24) and (3.25), and the corresponding PDE solutions are given in Supplementary Figures SM3 and SM4, respectively. The number of MGRIT iterations required to reach convergence on these problems is given in Table 1. The solution of coarse-grid linear systems involving the matrix $\mathcal{B}_{p+1}^{(t_n, m\delta t)}$ is approximated using 10 GMRES iterations with a zero initial guess. Fewer than 10 GMRES iterations can be used without impacting the results, but the focus of these particular tests is to determine the MGRIT convergence rate, independent of the cost of solving the coarse-grid linear systems. Note also for reasons relating to using GMRES inside MGRIT, we use a linear version of MGRIT, which we have implemented in XBraid (by default, XBraid uses the FAS framework); see Section SM2 for details.

Generally speaking, the convergence rates in Table 1 are fast for the two-level solution of hyperbolic problems. A general trend among these results is the convergence rate deteriorating with increasing discretization order, consistent with Figure 3. For the constant-wave-speed case, the iteration counts in Table 1 are essentially constant as the space-time mesh is refined. For the variable-wave-speed cases, there is some growth in iteration counts for the two high-order discretizations; at the same time, the iterations for these variable-wave-speed problems are typically smaller than those for the constant-wave-speed problem.

3.3.1. Scalable strategy for estimating coarse-grid departure points.

We now describe the strategy for estimating coarse-grid departure points that was used to generate the results in Table 1. Recall that in applying the coarse-grid

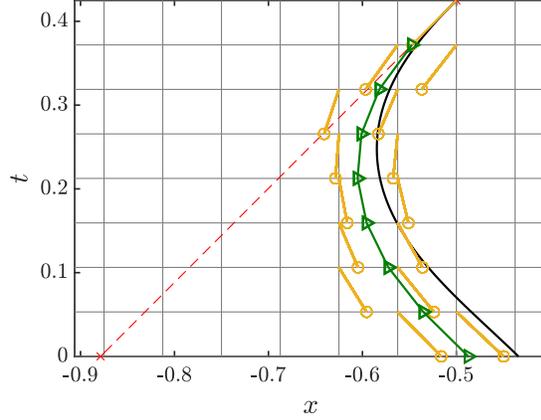


FIG. 4. Evolution of a local coarse-grid characteristic of advection problem (2.1) with wave-speed (3.25) using piecewise linear interpolation and backtracking of fine-grid characteristics obtained with an ERK method of accuracy $r = 1$. Note that only a subset of the spatial domain $x \in \Omega = (-1, 1)$ is shown to better highlight the detail of the characteristic. The black curve is the exact coarse-grid characteristic. The green curve is the coarse-grid characteristic approximated with the interpolation strategy (i.e., the triangle marker at time $t = k\delta t$ is $c_i^{(k)}$ from (3.27)). The gold lines are the fine-grid characteristics that are the nearest neighbors of the approximate coarse-grid characteristic (i.e., the left and right circle markers at time $t = k\delta t$ are, respectively, $f_W^{(k)}$ and $f_E^{(k)}$ from (3.27)). These fine-grid characteristics were determined by a single ERK step of size δt . The red dashed line is the coarse-grid characteristic approximated by a single ERK step of size $m\delta t = 8\delta t$.

semi-Lagrangian operator $\mathcal{S}_{p,r_*}^{(t_n, m\delta t)}$, we need to compute the values at time t_n of the local coarse-grid characteristics $\xi_i^{(t_n, m\delta t)}(t)$ that arrive at $(x, t) = (x_i, t_n + m\delta t)$. When time-stepping across the interval $t \in [t_n, t_n + m\delta t]$ on the fine grid with $\mathcal{S}_{p,r}^{(t_n + k\delta t, \delta t)}$, $k = 0, 1, \dots, m-1$, we map out the trajectories of the fine-grid characteristics $\xi_i^{(t_n + k\delta t, \delta t)}(t)$ over the fine-grid subintervals $t \in [t_n + k\delta t, t_n + (k+1)\delta t]$ (see the gold lines in Figure 4, with the circle markers representing the fine-grid departure points). In computing these fine-grid characteristics using an ERK method of order r , we in effect map out the vector field that dictates the flow of any local coarse-grid characteristic across the coarse space-time slab $(x, t) \in \Omega \times [t_n, t_n + m\delta t]$ with fine-grid-scale accuracy. The idea we propose now is to approximately propagate a coarse-grid characteristic through this space-time slab by recycling the fine-grid characteristic directions to guide its path in an interpolating manner, following the schematic shown in Figure 4.

For simplicity of notation, let us only consider the first coarse time interval $t \in [0, m\delta t]$. Let $c_i^{(k)}$, $k \in \{0, 1, \dots, m-1\}$, denote our approximation to the local coarse-grid characteristic $\xi_i^{(0, m\delta t)}(t)$ at time $t = k\delta t$, $c_i^{(k)} \approx \xi_i^{(0, m\delta t)}(k\delta t)$. For shorthand, denote departure points of the local fine-grid characteristics on this interval by $f_i^{(k)} \equiv \xi_i^{(k\delta t, \delta t)}(k\delta t)$, $k \in \{0, 1, \dots, m-1\}$.

Over the last fine-grid subinterval $t \in [(m-1)\delta t, m\delta t]$, the approximate coarse-grid characteristic is the same as the fine-grid characteristic, since they both arrive at $(x, t) = (x_i, m\delta t)$, and therefore they intersect the x axis at the same location,

$$(3.26) \quad c_i^{(m-1)} = f_i^{(m-1)}.$$

Using this as a final-time condition, the remaining intersection points of the coarse-

grid characteristic can be estimated by carrying out the following interpolating update in sequence

$$(3.27) \quad c_i^{(k)} = \frac{f_E^{(k)} - f_W^{(k)}}{h} \left(c_i^{(k+1)} - x_E \right) + f_E^{(k)}, \quad \text{for } k = m - 2, \dots, 1, 0,$$

where the integers E and W are such that x_E and x_W are the east- and west-neighboring mesh points of $c_i^{(k+1)}$, respectively. A schematic of this procedure is shown in [Figure 4](#). Upon completing the iteration (3.27), we have an approximation for the i th coarse-grid departure point, $c_i^{(0)} \approx \xi_i^{(0, m\delta t)}(0)$. The update formula (3.27) is based on piecewise linear nearest neighbor interpolation to estimate $c_i^{(k)}$. The example in [Figure 4](#) shows how this strategy has the potential to approximate coarse-grid departure points much more accurately than a single $m\delta t$ -sized step of the ERK scheme used for fine-grid characteristics.⁴

Recall that the motivation for the proposed strategy was to estimate coarse-grid departure points in a way that is less expensive than taking m ERK steps of size δt . However, since this linear interpolation and backtracking strategy requires taking $m - 1$ steps, it cannot, on the first coarse level, be significantly cheaper than using m steps of an ERK scheme.⁵ However, supposing that the linear interpolation strategy yields sufficiently accurate departure points, in the sense that its use does not lead to a strong deterioration of MGRIT convergence, then it has the significant advantage over stepping at the fine-grid-scale with an ERK scheme that it becomes cheaper on coarser levels in a way that makes the cost scalable to multiple levels. That is, say, for example, we have a three-level method in which we coarsen by m on each level. The linear interpolation strategy takes $\mathcal{O}(m)$ work per coarse time-step to estimate departure points on the first coarse level, but if it is then applied recursively on the coarsest level, it requires only $\mathcal{O}(m)$ work there to estimate a departure point. In general, if the strategy is applied recursively throughout a multilevel solver, it requires only $\mathcal{O}(m)$ work to estimate a departure point, independent of the level it occurs on. In contrast, if an ERK method is to be used to estimate departure points on coarse levels, by our previous arguments regarding the inaccuracy of taking large time-steps, it must do so by taking many small time-steps, with size of order the fine-grid time-step δt . That is, using an ERK method to estimate a coarse-grid departure point on a coarse level $\ell \in \mathbb{N}$ requires $\mathcal{O}(m^\ell)$ work.

Finally, we remark that the linear interpolation and backtracking strategy proposed here is more expensive from a memory perspective, since estimating departure points on a coarse level requires storing all departure points on the level above it.

3.4. Multilevel setting. In this section, we generalize the two-level, backward Euler coarse-grid operator (3.22) from the previous section so that it can be applied within a multilevel MGRIT algorithm. Let $\ell \in \mathbb{N}_0$ be the level index, and assume that the time-step size on level ℓ is $m^\ell \delta t$. To begin, we introduce the shorthand for

⁴In fact, if the wave-speed is spatially independent, then the strategy shown in [Figure 4](#) yields the same estimate for coarse-grid departure points as taking m steps of size δt with the fine-grid ERK scheme.

⁵Whether it is cheaper or not depends on the number of stages of the ERK scheme and the cost of evaluating the wave-speed. Recall that an s -stage ERK scheme requires s evaluations of the wave-speed per time-step. The linear interpolation strategy requires no evaluations of the wave-speed. In any event, the number of FLOPs for either strategy scales as $\mathcal{O}(m)$.

the following function on level $\ell \in \mathbb{N}$,

$$(3.28) \quad \varphi_{p+1}^{(t_n, m^\ell \delta t)} := (-1)^{p+1} \left(f_{p+1}(\varepsilon^{(t_n, m^\ell \delta t)}) - \sum_{k=0}^{m-1} f_{p+1}(\varepsilon^{(t_n + km^{\ell-1} \delta t, m^{\ell-1} \delta t)}) \right),$$

which generalizes the function $\varphi_{p+1}^{(t_n, m \delta t)}$ on level $\ell = 1$ defined in (3.17). Note that $\varphi_{p+1}^{(t_n, m^\ell \delta t)}$ approximates the coefficient vector appearing in the leading-order term of the difference between the level ℓ coarse-grid semi-Lagrangian operator, and the ideal coarse-grid operator defined by stepping across the same interval m times with the associated level $\ell - 1$ semi-Lagrangian operators, $\mathcal{S}_{p,\infty}^{(t_n, m^\ell \delta t)} - \prod_{k=0}^{m-1} \mathcal{S}_{p,\infty}^{(t_n + km^{\ell-1} \delta t, m^{\ell-1} \delta t)}$.

To develop a multilevel operator based on the backward Euler operator (3.22), it is first instructive to consider a three-level algorithm. Given the backward Euler operators $\Phi^{(t_n + km \delta t, m \delta t)} = \mathcal{B}_{p+1}^{(t_n + km \delta t, m \delta t)} \mathcal{S}_{p,\infty}^{(t_n + km \delta t, m \delta t)}$ on level $\ell = 1$ for $k \in \{0, \dots, m-1\}$, consider the associated ideal operator on level $\ell = 2$ and the following sequence of approximations to it,

$$(3.29)$$

$$(3.30) \quad \begin{aligned} \Phi_{\text{ideal}}^{(t_n, m^2 \delta t)} &= \prod_{k=0}^{m-1} \Phi^{(t_n + km \delta t, m \delta t)} = \prod_{k=0}^{m-1} \left(\mathcal{B}_{p+1}^{(t_n + km \delta t, m \delta t)} \mathcal{S}_{p,r_*}^{(t_n + km \delta t, m \delta t)} \right), \\ &\approx \left(\prod_{k=0}^{m-1} \mathcal{B}_{p+1}^{(t_n + km \delta t, m \delta t)} \right) \left(\prod_{k=0}^{m-1} \mathcal{S}_{p,r_*}^{(t_n + km \delta t, m \delta t)} \right), \end{aligned}$$

$$(3.31) \quad \approx \left(\prod_{k=0}^{m-1} \mathcal{B}_{p+1}^{(t_n + km \delta t, m \delta t)} \right) \left(\mathcal{B}_{p+1}^{(t_n, m^2 \delta t)} \mathcal{S}_{p,r_*}^{(t_n, m^2 \delta t)} \right),$$

$$(3.32) \quad \begin{aligned} &= \left(\prod_{k=0}^{m-1} \left[I - \text{diag} \left(\varphi_{p+1}^{(t_n + km \delta t, m \delta t)} \right) \mathcal{D}_{p+1} \right]^{-1} \right) \times \\ &\quad \left[I - \text{diag} \left(\varphi_{p+1}^{(t_n, m^2 \delta t)} \right) \mathcal{D}_{p+1} \right]^{-1} \mathcal{S}_{p,r_*}^{(t_n, m^2 \delta t)}, \end{aligned}$$

$$(3.33) \quad \approx \left[I - \text{diag} \left(\sum_{k=0}^{m-1} \varphi_{p+1}^{(t_n + km \delta t, m \delta t)} \right) \mathcal{D}_{p+1} \right]^{-1} \times$$

$$(3.34) \quad \left[I - \text{diag} \left(\varphi_{p+1}^{(t_n, m^2 \delta t)} \right) \mathcal{D}_{p+1} \right]^{-1} \mathcal{S}_{p,r_*}^{(t_n, m^2 \delta t)},$$

$$(3.34) \quad \approx \left[I - \text{diag} \left(\sum_{k=0}^{m-1} \varphi_{p+1}^{(t_n + km \delta t, m \delta t)} + \varphi_{p+1}^{(t_n, m^2 \delta t)} \right) \mathcal{D}_{p+1} \right]^{-1} \mathcal{S}_{p,r_*}^{(t_n, m^2 \delta t)}.$$

The approximation in (3.30) is that the backward Euler and semi-Lagrangian operators commute. Only in the case of spatially independent wave-speed functions do these two operators commute (since the diagonal matrices built from the various φ_{p+1} vectors are constant, and \mathcal{D}_{p+1} and the semi-Lagrangian operators are circulant). In any event, the forthcoming numerical results show this approximation is accurate enough to obtain fast MGRIT convergence when the wave-speed does depend on space.

In (3.31), m successive $m \delta t$ -sized semi-Lagrangian steps have been approximated using our existing two-level approximation. That is, the m steps are replaced by a single $m^2 \delta t$ -sized semi-Lagrangian step followed by a backward Euler step that approximately corrects for the lowest-order difference between their truncation errors.

The approximation in (3.33) is pulling the m backward Euler steps under the inverse, and keeping only the lowest-order terms in their product, recalling $\mathcal{D}_{p+1}\mathbf{v} = \mathcal{O}(h^{p+1})$. This approximation can be understood as a Taylor series interpretation of the standard rediscrization approach typically employed in MGRIT for backward Euler discretizations, in which m backward Euler steps are approximated on the coarse level with a single backward Euler step using a time-step size that is m times larger.

Finally, (3.34) arises from placing the two backward Euler matrices under a single inverse, taking their product, and then truncating the highest-order term, which is proportional to $\mathcal{D}_{p+1}\mathcal{D}_{p+1}$. Notice that (3.34) has the same structure as the operator (3.22) on level $\ell = 1$ proposed for the two-level algorithm, since it is a semi-Lagrangian step followed by a backward Euler correction. Based on this, we propose the following time-stepping operators on level $\ell > 0$ for evolving solutions from $t_n \rightarrow t_n + m^\ell \delta t$,

$$(3.35) \quad \Phi^{(t_n, m^\ell \delta t)} = \left[I - \text{diag} \left(\boldsymbol{\sigma}_{p+1}^{(t_n, m^\ell \delta t)} \right) \mathcal{D}_{p+1} \right]^{-1} \mathcal{S}_{p, r_*}^{(t_n, m^\ell \delta t)}, \quad \ell \in \mathbb{N},$$

in which the coefficient vector is defined recursively by

$$(3.36) \quad \boldsymbol{\sigma}_{p+1}^{(t_n, m^\ell \delta t)} = \begin{cases} \boldsymbol{\varphi}_{p+1}^{(t_n, m^\ell \delta t)}, & \ell = 1, \\ \sum_{k=0}^{m-1} \boldsymbol{\sigma}_{p+1}^{(t_n + km^{\ell-1} \delta t, m^{\ell-1} \delta t)} + \boldsymbol{\varphi}_{p+1}^{(t_n, m^\ell \delta t)}, & \ell > 1. \end{cases}$$

We now present results of our numerical tests using the coarse-grid operator (3.35). In these tests, we use a constant coarsening factor of m on all levels, and continue to coarsen until doing so would result in fewer than two points in time. To locate departure points on coarse levels, the linear interpolation and backtracking strategy of Section 3.3.1 is employed recursively. Furthermore, we now slightly change our strategy for approximately inverting the backward Euler matrix on coarse levels. Specifically, for each linear system, we iterate GMRES until the norm of the relative residual decreases below 10^{-2} or the number of iterations reaches 10.

		$\alpha = (3.23)$			$\alpha = (3.24)$			$\alpha = (3.25)$		
		m			m			m		
p, r	$n_x \times n_t$	4	8	16	4	8	16	4	8	16
1, 1	$2^8 \times 2^{10}$	14	12	11	12	11	11	11	11	12
	$2^{10} \times 2^{12}$	14	12	11	12	11	11	12	12	12
	$2^{12} \times 2^{14}$	14	13	11	14	11	11	14	13	13
3, 3	$2^8 \times 2^{10}$	23	17	15	15	13	12	15	15	15
	$2^{10} \times 2^{12}$	23	17	15	16	14	13	16	16	16
	$2^{12} \times 2^{14}$	23	18	16	21	15	13	20	19	19
5, 5	$2^8 \times 2^{10}$	32	22	18	19	16	14	17	19	17
	$2^{10} \times 2^{12}$	34	24	20	19	16	15	19	19	20
	$2^{12} \times 2^{14}$	34	24	21	27	18	15	23	24	25

TABLE 2

Number of MGRIT V-cycles to converge. The PDE is (2.1) with wave-speed $\alpha(x, t)$ indicated in the top row of the table. The fine-grid operator is $\mathcal{S}_{p, r}^{(t_n, \delta t)}$, and the coarse-grid operator is the dissipatively corrected operator (3.35). Departure points on coarse levels are located by recursively applying the strategy from Section 3.3.1. A coarsening factor of m is used on all levels.

The MGRIT V-cycle iteration counts for these tests are given in [Table 2](#). Notice that many of the iteration counts are almost identical to those for the two-level tests given in [Table 1](#). We therefore conclude that the multilevel coarse-grid operator [\(3.35\)](#) performs as well as one could anticipate given the performance of the two-level operator that it generalizes. We note that this is novel because it is not uncommon to see multigrid iterations strongly increase for hyperbolic problems when transitioning from two to many levels [\[25, 46, 47, 23\]](#). These promising results also indicate that our linear interpolation and backtracking strategy of [Section 3.3.1](#) for approximating coarse-grid departure points does so with a degree of accuracy that does not hamper multilevel MGRIT convergence, even on much coarser levels.

4. Two spatial dimensions. We now extend the coarse-grid operator from the previous section to advection problems in two spatial dimensions. [Section 4.1](#) discusses the semi-Lagrangian discretization, the coarse-grid operator is presented in [Section 4.2](#), and [Section 4.3](#) presents numerical results.

4.1. Semi-Lagrangian discretization. We now consider semi-Lagrangian discretizations of two-dimensional advection problems of the form

$$(4.1) \quad \frac{\partial u}{\partial t} + \alpha(x, y, t) \frac{\partial u}{\partial x} + \beta(x, y, t) \frac{\partial u}{\partial y} = 0, \quad (x, y, t) \in \Omega \times (0, T],$$

with initial condition $u(x, y, 0) = u_0(x, y)$, spatial domain $\Omega \subset \mathbb{R}^2$, and solution u subject to periodic boundary conditions on $\partial\Omega$. Specifically, our numerical tests for this two-dimensional problem will use the initial condition $u(x, y, 0) = \sin^2 \left[\frac{\pi}{2}(x - 1) \right] \sin^2 \left[\frac{\pi}{2}(y - 1) \right]$, and the spatial domain $\Omega = (-1, 1)^2$. The semi-Lagrangian discretizations we consider of [\(4.1\)](#) are a straightforward generalization of those described in [Section 2.1](#) for the one-dimensional problem.

We define a discrete mesh on Ω as the tensor product of one-dimensional meshes in the x - and y -directions, respectively, both of which we assume are composed of n_x points equispaced by a distance of h . Let $(x, y, t) = (\xi(t), \eta(t), t)$ denote a characteristic of [\(4.1\)](#), then the Lagrangian formulation of [\(4.1\)](#) reads

$$(4.2) \quad \frac{d}{dt} \xi(t) = \alpha(\xi(t), \eta(t), t), \quad \frac{d}{dt} \eta(t) = \beta(\xi(t), \eta(t), t), \quad \frac{d}{dt} u(\xi(t), \eta(t), t) = 0.$$

Define $(\xi_{ij}^{(t_n, \delta t)}(t), \eta_{ij}^{(t_n, \delta t)}(t))$ as the local characteristic that passes through the *arrival point* $(x, y, t) = (\xi_{ij}^{(t_n, \delta t)}(t_{n+1}), \eta_{ij}^{(t_n, \delta t)}(t_{n+1}), t_{n+1})$. Then, the associated *departure point* $(x, y, t) = (\xi_{ij}^{(t_n, \delta t)}(t_n), \eta_{ij}^{(t_n, \delta t)}(t_n), t_n)$ is given by the solution at time $t = t_n$ of the following final-value problem that holds over $t \in [t_n, t_{n+1})$

$$(4.3) \quad \frac{d}{dt} \xi_{ij}^{(t_n, \delta t)}(t) = \alpha(\xi_{ij}^{(t_n, \delta t)}(t), \eta_{ij}^{(t_n, \delta t)}(t), t), \quad \xi_{ij}^{(t_n, \delta t)}(t_{n+1}) = x_i,$$

$$(4.4) \quad \frac{d}{dt} \eta_{ij}^{(t_n, \delta t)}(t) = \beta(\xi_{ij}^{(t_n, \delta t)}(t), \eta_{ij}^{(t_n, \delta t)}(t), t), \quad \eta_{ij}^{(t_n, \delta t)}(t_{n+1}) = y_j.$$

Upon (approximately) locating the departure point of the local characteristic, the solution is estimated at it via two-dimensional polynomial interpolation through its nearest neighboring mesh points. Generalizing what we did in the one-dimensional case, let $(x, y) = (x_{ij}^{(t_n, \delta t)}, y_{ij}^{(t_n, \delta t)})$ be the mesh point immediately to the north-east of the departure point $(x, y) = (\xi_{ij}^{(t_n, \delta t)}(t_n), \eta_{ij}^{(t_n, \delta t)}(t_n))$. Then, decompose the x -coordinate of the departure point as $\xi_{ij}^{(t_n, \delta t)}(t_n) \equiv x_{ij}^{(t_n, \delta t)} - h\varepsilon_{ij}^{(t_n, \delta t)}$, $\varepsilon_{ij}^{(t_n, \delta t)} \in [0, 1)$, and the y -coordinate as $\eta_{ij}^{(t_n, \delta t)}(t_n) \equiv y_{ij}^{(t_n, \delta t)} - h\nu_{ij}^{(t_n, \delta t)}$, $\nu_{ij}^{(t_n, \delta t)} \in [0, 1)$.

The two-dimensional interpolating polynomial is then constructed through a tensor product of a one-dimensional interpolation in the x -direction and a one-dimensional interpolation in the y -direction. See [14, pp. 61–62] for further details.

4.2. Coarse-grid operator. We now generalize the one-dimensional considerations of Section 3 to develop a coarse-grid operator for the two-dimensional problem (4.1). In the following, the matrix $\mathcal{D}_{p+1} \in \mathbb{R}^{n_x \times n_x}$ is defined as in (3.1). Supposing spatial degrees-of-freedom are ordered row-wise lexicographically, applying $h^{-(p+1)}(I_{n_x} \otimes \mathcal{D}_{p+1})$ or $h^{-(p+1)}(\mathcal{D}_{p+1} \otimes I_{n_x})$ to a periodic grid vector gives an approximation to its $p+1$ st partial derivative with respect to x or y , respectively.

LEMMA 4.1 (Semi-Lagrangian truncation error for $r = \infty$). *Let \mathcal{D}_{p+1} be as in (3.1). Suppose that the solution $u(x, y, t)$ of (4.1) is at least $p+1+s$ times continuously differentiable with respect to x and y , and that the wave-speed in (4.1) is independent of space, $(\alpha(x, y, t), \beta(x, y, t)) \equiv (\alpha(t), \beta(t))$. Define $\mathbf{u}(t) \in \mathbb{R}^{n_x^2}$ as the vector composed of the PDE solution sampled in space at the mesh points and at time t , and let $\mathcal{S}_{p,\infty}^{(t_n, \delta t)}$ be the semi-Lagrangian discretization of (4.1) that exactly locates departure points at time t_n . Then, the local truncation error of this discretization can be expressed as*

$$(4.5) \quad \begin{aligned} \mathbf{u}(t_{n+1}) - \mathcal{S}_{p,\infty}^{(t_n, \delta t)} \mathbf{u}(t_n) &= (-h)^{p+1} \left[\text{diag} \left(f_{p+1}(\boldsymbol{\varepsilon}^{(t_n, \delta t)}) \right) \frac{(I_{n_x} \otimes \mathcal{D}_{p+1})}{h^{p+1}} \right. \\ &\quad \left. + \text{diag} \left(f_{p+1}(\boldsymbol{\nu}^{(t_n, \delta t)}) \right) \frac{(\mathcal{D}_{p+1} \otimes I_{n_x})}{h^{p+1}} \right] \mathbf{u}(t_{n+1}) + \mathcal{O}(h^{p+2}), \end{aligned}$$

in which the entries of the vectors $\boldsymbol{\varepsilon}^{(t_n, \delta t)}, \boldsymbol{\nu}^{(t_n, \delta t)} \in \mathbb{R}^{n_x^2}$ associated with the mesh point $(x, y) = (x_i, y_j)$ are $\varepsilon_{ij}^{(t_n, \delta t)}$, and $\nu_{ij}^{(t_n, \delta t)}$, respectively. The polynomial f_{p+1} is given in (3.3).

Furthermore, the ideal coarse-grid operator defined by time-stepping across $t \in [t_n, t_n + m\delta t]$ with the m fine-grid operators $\mathcal{S}_{p,\infty}^{(t_n + k\delta t, \delta t)}$, $k = 0, \dots, m-1$, has a local truncation error given by

$$(4.6) \quad \begin{aligned} \mathbf{u}(t_{n+m}) - \left[\prod_{k=0}^{m-1} \mathcal{S}_{p,\infty}^{(t_n + k\delta t, \delta t)} \right] \mathbf{u}(t_n) &= (-h)^{p+1} \left[\text{diag} \left(\sum_{k=0}^{m-1} f_{p+1}(\boldsymbol{\varepsilon}^{(t_n + k\delta t, \delta t)}) \right) \times \right. \\ &\quad \left. \frac{(I_{n_x} \otimes \mathcal{D}_{p+1})}{h^{p+1}} + \text{diag} \left(\sum_{k=0}^{m-1} f_{p+1}(\boldsymbol{\nu}^{(t_n + k\delta t, \delta t)}) \right) \frac{(\mathcal{D}_{p+1} \otimes I_{n_x})}{h^{p+1}} \right] \mathbf{u}(t_{n+m}) + \mathcal{O}(h^{p+2}). \end{aligned}$$

Proof. We omit details of these proofs since they follow analogously to those for the one-dimensional cases given in Lemma 3.1 and Corollary 3.2. The one caveat is that an error estimate for two-dimensional polynomial interpolation needs to be applied (the required estimate is given as [27, Lem. 4.11]). \square

Generalizing the one-dimensional coarse-grid operator from Section 3.2, we propose the following coarse-grid operator for semi-Lagrangian discretizations $\mathcal{S}_{p,r}^{(t_n, \delta t)}$ of

the two-dimensional advection problem (4.1)

(4.7)

$$\begin{aligned} \Phi_{p,r_*}^{(t_n,\delta t)} &= \left[I_{n_x^2} - \text{diag} \left(f_{p+1}(\boldsymbol{\varepsilon}^{(t_n,m\delta t)}) - \sum_{k=0}^{m-1} f_{p+1}(\boldsymbol{\varepsilon}^{(t_n+k\delta t,\delta t)}) \right) (I_{n_x} \otimes \mathcal{D}_{p+1}) \right. \\ &\quad \left. - \text{diag} \left(f_{p+1}(\boldsymbol{\nu}^{(t_n,m\delta t)}) - \sum_{k=0}^{m-1} f_{p+1}(\boldsymbol{\nu}^{(t_n+k\delta t,\delta t)}) \right) (\mathcal{D}_{p+1} \otimes I_{n_x}) \right]^{-1} \mathcal{S}_{p,r_*}^{(t_n,m\delta t)}. \end{aligned}$$

4.3. Numerical results. We now present numerical results for solving the two-dimensional advection equation (4.1) with the constant wave-speed

$$(4.8) \quad (\alpha(x, y, t), \beta(x, y, t)) = (1, 1),$$

as well as the variable wave-speed

$$(4.9) \quad (\alpha(x, y, t), \beta(x, y, t)) = \left(\sin^2(\pi y) \cos\left(\frac{2\pi t}{3.4}\right), -\cos^2(\pi x) \cos\left(\frac{2\pi t}{3.4}\right) \right).$$

Plots of the velocity field and the solution of PDE (4.1) associated with the wave-speed (4.9) are given in Figures SM6 and SM7, respectively. Note that the PDE solution associated with (4.9) is periodic in time, with period 3.4.

The numerical tests use a fine-grid time-step size of $\delta t = 0.85h$. In Table 3, we report the number of MGRIT V-cycles to converge using the dissipatively corrected operator (4.7). On coarse levels $\ell > 1$, we use a straightforward generalization of the

p, r	$n_x^2 \times n_t$	$(\alpha, \beta) = (4.8)$			$(\alpha, \beta) = (4.9)$		
		m			m		
		4	8	16	4	8	16
1, 1	$(2^6)^2 \times 2^{10}$	14 (50)	12 (45)	11 (27)	12 (68)	11 (44)	12 (27)
	$(2^7)^2 \times 2^{11}$	14 (86)	12 (81)	11 (46)	12 (157)	12 (80)	11 (46)
	$(2^8)^2 \times 2^{12}$	14	12	11	12	12	12
	$(2^9)^2 \times 2^{13}$	14	12	11	14	13	13
3, 3	$(2^6)^2 \times 2^{10}$	22 (57)	17 (47)	15 (27)	16 (70)	15 (44)	15 (27)
	$(2^7)^2 \times 2^{11}$	23 (98)	17 (84)	15 (47)	15 (170)	16 (81)	16 (46)
	$(2^8)^2 \times 2^{12}$	23	17	15	16	16	17
	$(2^9)^2 \times 2^{13}$	23	17	16	19	18	17
5, 5	$(2^6)^2 \times 2^{10}$	32 (63)	22 (48)	18 (27)	19 (69)	18 (44)	18 (27)
	$(2^7)^2 \times 2^{11}$	33 (110)	23 (88)	20 (48)	19 (169)	19 (80)	20 (46)
	$(2^8)^2 \times 2^{12}$	34	24	20	19	21	22
	$(2^9)^2 \times 2^{13}$	34	24	21	24	23	23

TABLE 3

Number of MGRIT V-cycles to reach convergence on the two-dimensional advection problem (4.1), with wave-speed indicated in the top row of the table. The fine-grid discretization is $\mathcal{S}_{p,r}^{(t_n,\delta t)}$, and the coarse-grid operator is the dissipatively corrected semi-Lagrangian operator (4.7). For the two smallest mesh resolutions $(n_x^2, n_t) = ((2^6)^2 \times 2^{10}, (2^6)^2 \times 2^{11})$, iterations reported in parenthesis corresponds to using, instead, the rediscrretized semi-Lagrangian operator $\mathcal{S}_{p,r}^{(t_n,m^\ell \delta t)}$ on coarse levels.

level $\ell = 1$ operator (4.7) by following the development of the multilevel operator in Section 3.4 for the one-dimensional case. Solutions of coarse-grid linear systems are approximated by iterating GMRES until either the relative residual norm is 10^{-2} or smaller, or the number of iterations reaches 10. The coarse-grid semi-Lagrangian operators $\mathcal{S}_{p,r_*}^{(t_n, m^\ell \delta t)}$ used in (4.7) estimate departure points using a generalization of the one-dimensional backtracking and linear interpolation strategy from Section 3.3.1; see Section SM6 for details. Finally, for the two smallest mesh resolutions, Table 3 also includes iteration counts for the naive choice of simply rediscrctizing the semi-Lagrangian discretization; in these rediscrctization experiments, departure points on a coarse-level $\ell > 0$ are estimated accurately using m^ℓ ERK steps of size δt .

Iteration counts in Table 3 for the rediscrctized coarse-grid operator are large, and grow strongly as the mesh is refined. Therefore, as for the one-dimensional case, we conclude that simply rediscrctizing the semi-Lagrangian coarse-grid operator does not lead to a robust MGRIT solver for our two-dimensional model problem. In contrast, the proposed modified semi-Lagrangian coarse-grid operator (4.7) yields much smaller iteration counts, that seem nearly constant as the mesh is refined. These sequential numerical results highlight the potential our proposed coarse-grid operator (4.7) has for parallel-in-time simulations of advection problems.

Remark 4.2 (Potential non-robustness for spatially varying wave-speeds). In a small number of test problems with spatially varying wave-speeds, we have observed less favorable MGRIT convergence when using the proposed coarse-grid operator. While we have sometimes also encountered these issues in one dimension, they seem most pronounced in two dimensions, and they appear related to the time-step size on some coarse level not being sufficiently small. Recall that we conjectured in Remark 3.4 that in one dimension, the error estimates for spatially independent wave-speed functions carry over to the spatially variable case up to terms of size $\mathcal{O}(h^{p+1}\delta t)$. We suspect that the robustness issues we describe above are a consequence of such additional $\mathcal{O}(h^{p+1}\delta t)$ terms that are not captured by the proposed coarse-grid operator. In particular, if one coarsens down to $\delta t = \mathcal{O}(1)$ on some coarse level, then such $\mathcal{O}(h^{p+1}\delta t)$ terms may no longer be small relative to the $\mathcal{O}(h^{p+1})$ terms that are included in the proposed coarse-grid operator. It remains on-going work to further investigate and address this issue, potentially by deriving estimates that explicitly include the $\mathcal{O}(h^{p+1}\delta t)$ terms and then incorporating them into the coarse-grid operator. Regardless, when δt is sufficiently small on all levels, as is the case for all the reported results, these issues do not occur.

5. Parallel results. In this section we present parallel strong-scaling results for MGRIT V-cycles using the coarse-grid operators developed in earlier sections. Since we want to demonstrate the efficacy of our proposed coarse-grid operator, we consider tests that use parallelism in time only, and we leave a space-time parallel implementation to future work. The results were generated on Ruby, a Linux cluster at Lawrence Livermore National Laboratory consisting of 1,480 compute nodes, with 56 Intel Xeon CLX-8276L cores per node. The results used the following node configurations: (number of MPI tasks, number of nodes) = ((1, 1), (4, 2), (16, 2), (64, 3), (256, 10), (1024, 37)). Tests using four and 16 MPI tasks were run across two nodes since we observed this was faster than on a single node. The number of nodes used in the larger tests was chosen to ensure the number of MPI tasks per node did not exceed 28 (the number of physical cores per node).

We consider strong-scaling studies for both the one- and two-dimensional advection problems (2.1) and (4.1), respectively. The tests use the largest problem sizes

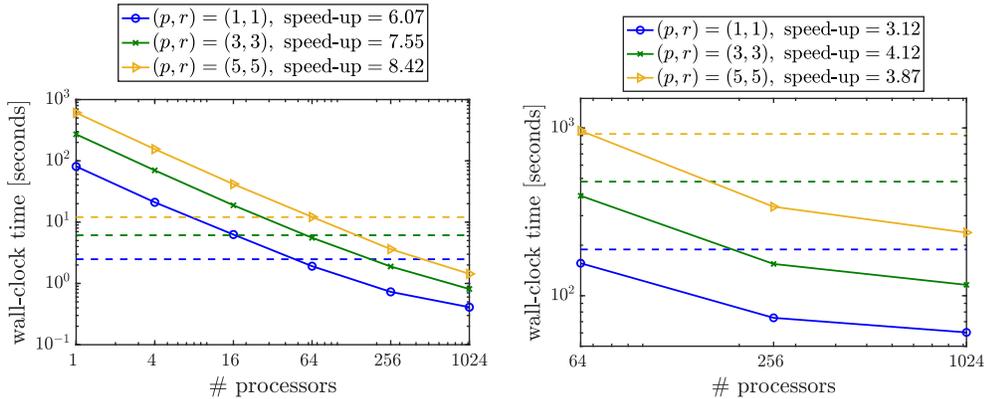


FIG. 5. *Strong parallel scaling using time-only parallelism: Runtimes of MGRIT V-cycles with an aggressive coarsening factor of $m = 16$ on the first level followed by $m = 4$ on all coarser levels. The semi-Lagrangian discretizations are of orders $(p, r) = ((1, 1), (3, 3), (5, 5))$. Left: The one-dimensional PDE (2.1) with wave-speed (3.25), on space-time mesh having $n_x \times n_t = 2^{12} \times 2^{14}$ points. Right: The two-dimensional PDE (4.1) with wave-speed (4.9), on a space-time mesh having $n_x^2 \times n_t = (2^9)^2 \times 2^{13}$ points. Dashed lines represent runtimes of time-stepping on one processor. Speed-ups obtained using 1024 processors are listed in the legends.*

considered previously of $n_x \times n_t = 2^{12} \times 2^{14}$, and $n_x^2 \times n_t = (2^9)^2 \times 2^{13}$ for the one- and two-dimensional problems, respectively (see Tables 2 and 3). For the one-dimensional problem, we have done scaling studies using coarsening factors $m = 4, 8, 16$, as well as an aggressive coarsening strategy using $m = 16$ on the first level and $m = 4$ on all other levels. For this problem, the aggressive coarsening strategy yielded the largest speed-ups, so we report results for this coarsening strategy only. Based on this finding we have adopted this aggressive coarsening strategy for our two-dimensional tests.

Runtimes as a function of processor count are shown in Figure 5. For the two-dimensional problem, a minimum of 64 processors is used since runtimes on fewer processors would have been too large. For both one- and two-dimensional problems, the cross-over point at which MGRIT is faster than time-stepping is around 64 processors. Notice that the speed-ups we achieve on 1024 processors for the two-dimensional problem are roughly a factor of two smaller than for the one-dimensional problem; this is perhaps unsurprising since half as many points in time are used, thus cutting the potential for speed-ups roughly in half. Additional one-dimensional tests solving to discretization error rather than reducing the ℓ^2 -norm of the space-time residual by 10 orders of magnitude are given in Section SM4.

Making fair comparisons between speed-ups obtained on different problems using different setups is difficult; however, it is fair to say that when using a similar number of processors, the speed-ups we obtain here are smaller than those achieved for MGRIT applied to diffusion-dominated problems, such as the heat equation [15, 17]. This is likely due to the increased complexity of our coarse-grid operator relative to rediscrretization (as is typically used for diffusion-dominated problems), and also the somewhat higher iteration counts we require to reach convergence. Nonetheless, the speed-ups we report are quite strong relative to speed-ups reported elsewhere in the literature for hyperbolic problems, especially since we also consider high-order discretizations [31, 25]. Furthermore, the speed-ups we obtain for the one-dimensional problem are somewhat comparable to those in [10] that used optimized coarse-grid operators for constant-wave-speed advection problems; the methods from [10], however,

are impractical for real computations and do not extend to variable wave-speeds.

6. Conclusions. Robust and efficient parallel-in-time integration of advection-dominated PDEs using the iterative multigrid-in-time method MGRIT, and the closely related Parareal method, is notoriously difficult. *Rediscretizing* the fine-grid problem on coarse grids often results in MGRIT diverging on advection-dominated PDEs, despite the same technique typically leading to excellent convergence for diffusion-dominated problems. To date, no practical coarse-grid operators have been proposed for these algorithms that are capable of providing speed-up over sequential time-stepping, even for the simplest case of constant-wave-speed linear advection.

We have considered a specific class of semi-Lagrangian discretizations for linear advection problems, including those in two spatial dimensions, and with variable wave-speeds. For these problems, we have proposed a novel modified semi-Lagrangian coarse-grid operator that, for the first time, leads to fast and robust MGRIT convergence. Parallel results show the coarse-grid operator can provide speed-up over sequential time-stepping, including for high-order discretizations. The proposed coarse-grid operator applies a semi-Lagrangian discretization followed by a correction, designed so that the truncation error of this composition approximately matches that of the so-called ideal coarse-grid operator. For the problems considered here, the correction adds dissipation to the coarse-grid semi-Lagrangian discretization via solving a linear system resembling a backward Euler time discretization of a heat-like equation.

The ideas presented here open many directions for future research, including adapting our approach to dispersive semi-Lagrangian schemes, since our current approach is only effective for dissipative semi-Lagrangian schemes. A key component of our on-going research is the extension of the ideas presented here to enable the robust MGRIT solution of more classical method-of-lines discretizations of hyperbolic PDEs; that is, coupling a fine-grid method-of-lines discretization with a coarse-grid operator that is semi-Lagrangian in nature with a truncation error perturbation. For constant-wave-speed problems, we have promising initial results for discretizations considered previously in [10], which use explicit and implicit Runge-Kutta in time, and finite differences in space. Another direction of future work is to extend these ideas to nonlinear conservation laws.

REFERENCES

- [1] G. BAL AND Y. MADAY, *A “parareal” time discretization for non-linear PDE’s with application to the pricing of an american put*, in Lecture Notes in Computational Science and Engineering, Springer Berlin Heidelberg, 2002, pp. 189–202.
- [2] R. E. BANK, J. W. L. WAN, AND Z. QU, *Kernel preserving multigrid methods for convection-diffusion equations*, SIAM J. Matrix Anal. Appl., 27 (2006), pp. 1150–1171.
- [3] A. BRANDT, *Multigrid solvers for non-elliptic and singular-perturbation steady-state problems*. The Weizmann Institute of Science, Rehovot, Israel. (unpublished), 1981.
- [4] A. BRANDT AND I. YAVNEH, *Accelerated multigrid convergence and high-reynolds recirculating flows*, SIAM J. Sci. Comput., 14 (1993), pp. 607–626.
- [5] X. CAI, S. BOSCARINO, AND J.-M. QIU, *High order semi-lagrangian discontinuous galerkin method coupled with runge-kutta exponential integrators for nonlinear vlasov dynamics*, J. Comput. Phys., 427 (2021), p. 110036.
- [6] X. CAI, W. GUO, AND J.-M. QIU, *A high order conservative semi-Lagrangian discontinuous Galerkin method for two-dimensional transport simulations*, J. Sci. Comput., 73 (2017), pp. 514–542.
- [7] F. CHEN, J. S. HESTHAVEN, AND X. ZHU, *On the use of reduced basis methods to accelerate and stabilize the parareal method*, in Reduced Order Methods for modeling and computational reduction, Springer, 2014, pp. 187–214.
- [8] X. DAI AND Y. MADAY, *Stable parareal in time method for first-and second-order hyperbolic*

- systems*, SIAM J. Sci. Comput., 35 (2013), pp. A52–A78.
- [9] P. J. DAVIS, *Interpolation and approximation*, Dover Publications, INC., New York, 1975.
- [10] H. DE STERCK, R. D. FALGOUT, S. FRIEDHOFF, O. A. KRZYSIK, AND S. P. MACLACHLAN, *Optimizing multigrid reduction-in-time and Parareal coarse-grid operators for linear advection*, Numer. Linear Algebra Appl., 28 (2021).
- [11] H. DE STERCK, S. FRIEDHOFF, A. J. M. HOWSE, AND S. P. MACLACHLAN, *Convergence analysis for parallel-in-time solution of hyperbolic systems*, Numer. Linear Algebra Appl., 27 (2020), p. e2271.
- [12] V. A. DOBREV, T. KOLEV, N. A. PETERSSON, AND J. B. SCHRODER, *Two-level convergence theory for multigrid reduction in time (MGRIT)*, SIAM J. Sci. Comput., 39 (2017), pp. S501–S527.
- [13] D. R. DURRAN, *Numerical Methods for Fluid Dynamics*, Springer New York, second ed., 2010.
- [14] M. FALCONE AND R. FERRETTI, *Semi-Lagrangian Approximation Schemes for Linear and Hamilton Jacobi Equations*, CAMBRIDGE, Feb. 2014.
- [15] R. D. FALGOUT, S. FRIEDHOFF, T. V. KOLEV, S. P. MACLACHLAN, AND J. B. SCHRODER, *Parallel time integration with multigrid*, SIAM J. Sci. Comput., 36 (2014), pp. C635–C661.
- [16] R. D. FALGOUT, M. LECOUEZ, AND C. S. WOODWARD, *A parallel-in-time algorithm for variable step multistep methods*, J. Comput. Sci., 37 (2019), p. 101029.
- [17] R. D. FALGOUT, T. A. MANTEUFFEL, B. O’NEILL, AND J. B. SCHRODER, *Multigrid reduction in time for nonlinear parabolic problems: A case study*, SIAM J. Sci. Comput., 39 (2017), pp. S298–S322.
- [18] M. J. GANDER, *Analysis of the parareal algorithm applied to hyperbolic problems using characteristics*, Soc. Esp. Mat. Apl., 42 (2008), pp. 21–35.
- [19] M. J. GANDER, *50 years of time parallel time integration*, in Contrib. Math. Comput. Sci., Springer International Publishing, 2015, pp. 69–113.
- [20] M. J. GANDER, L. HALPERN, J. RANNOU, AND J. RYAN, *A direct time parallel solver by diagonalization for the wave equation*, SIAM J. Sci. Comput., 41 (2019), pp. A220–A245.
- [21] M. J. GANDER, F. KWOK, AND H. ZHANG, *Multigrid interpretations of the parareal algorithm leading to an overlapping variant and MGRIT*, Comput. Vis. Sci., 19 (2018), pp. 59–74.
- [22] M. J. GANDER AND S. VANDEWALLE, *Analysis of the parareal time-parallel time-integration method*, SIAM J. Sci. Comput., 29 (2007), pp. 556–578.
- [23] A. HESSENTHALER, D. NORDSLETTEN, O. RÖHRLE, J. B. SCHRODER, AND R. D. FALGOUT, *Convergence of the multigrid reduction in time algorithm for the linear elasticity equations*, Numer. Linear Algebra Appl., 25 (2018), p. e2155.
- [24] A. HOWSE, *Nonlinear Preconditioning Methods for Optimization and Parallel-In-Time Methods for 1D Scalar Hyperbolic Partial Differential Equations*, PhD thesis, University of Waterloo, Waterloo, Canada, 2017.
- [25] A. J. M. HOWSE, H. DE STERCK, R. D. FALGOUT, S. MACLACHLAN, AND J. SCHRODER, *Parallel-in-time multigrid with adaptive spatial coarsening for the linear advection and inviscid Burgers equations*, SIAM J. Sci. Comput., 41 (2019), pp. A538–A565.
- [26] C.-S. HUANG, T. ARBOGAST, AND C.-H. HUNG, *A semi-Lagrangian finite difference WENO scheme for scalar nonlinear conservation laws*, J. Comput. Phys., 322 (2016), pp. 559–585.
- [27] O. A. KRZYSIK, *Multilevel parallel-in-time methods for advection-dominated PDEs*, Monash University, (2021), <https://doi.org/10.26180/17108537.v1>.
- [28] J.-L. LIONS, Y. MADAY, AND G. TURINICI, *Résolution d’edp par un schéma en temps pararéel*, C. R. Acad. Sci-Series I-Mathematics, 332 (2001), pp. 661–668.
- [29] J. LIU AND S.-L. WU, *A fast block α -circulant preconditioner for all-at-once systems from wave equations*, SIAM J. Matrix Anal. Appl., 41 (2020), pp. 1912–1943.
- [30] G. MENGALDO, A. WYSZOGRODZKI, M. DIAMANTAKIS, S.-J. LOCK, F. X. GIRALDO, AND N. P. WEDI, *Current and emerging time-integration strategies in global numerical weather and climate prediction*, Arch. Comput. Methods Eng., 26 (2018), pp. 663–684.
- [31] A. S. NIELSEN, G. BRUNNER, AND J. S. HESTHAVEN, *Communication-aware adaptive Parareal with application to a nonlinear hyperbolic system of partial differential equations*, J. Comput. Phys., 371 (2018), pp. 483–505.
- [32] B. W. ONG AND J. B. SCHRODER, *Applications of time parallelization*, Comput. Vis. Sci., 23 (2020).
- [33] J.-M. QIU AND C.-W. SHU, *Conservative high order semi-lagrangian finite difference WENO methods for advection in incompressible flow*, J. Comput. Phys., 230 (2011), pp. 863–889.
- [34] D. RUPRECHT, *Wave propagation characteristics of Parareal*, Comput. Vis. Sci., 19 (2018), pp. 1–17.
- [35] A. SCHMITT, M. SCHREIBER, P. PEIXOTO, AND M. SCHÄFER, *A numerical study of a semi-Lagrangian Parareal method applied to the viscous Burgers equation*, Comput. Vis. Sci.,

- 19 (2018), pp. 45–57.
- [36] J. B. SCHRODER, *On the use of artificial dissipation for hyperbolic problems and multigrid reduction in time (MGRIT)*. LLNL Tech Report LLNL-TR-750825, 2018.
 - [37] A. A. SIVAS, B. S. SOUTHWORTH, AND S. RHEBERGEN, *AIR algebraic multigrid for a space-time hybridizable discontinuous galerkin discretization of advection(-diffusion)*, SIAM Journal on Scientific Computing, 43 (2021), pp. A3393–A3416.
 - [38] P. K. SMOLARKIEWICZ AND J. A. PUDYKIEWICZ, *A class of semi-lagrangian approximations for fluids*, J. Atmospheric Sci., 49 (1992), pp. 2082–2096.
 - [39] B. S. SOUTHWORTH, *Necessary conditions and tight two-level convergence bounds for parareal and multigrid reduction in time*, SIAM J. Matrix Anal. Appl., 40 (2019), pp. 564–608.
 - [40] A. STANFORTH AND J. CÔTÉ, *Semi-lagrangian integration schemes for atmospheric models—a review*, Mon. Wea. Rev., 119 (1991), pp. 2206–2223.
 - [41] J. STEINER, D. RUPRECHT, R. SPECK, AND R. KRAUSE, *Convergence of Parareal for the Navier-Stokes equations depending on the Reynolds number*, in Numerical Mathematics and Advanced Applications-ENUMATH 2013, Springer, 2015, pp. 195–202.
 - [42] U. TROTTEBERG, C. W. OOSTERLEE, AND A. SCHULLER, *Multigrid*, Academic press, 2001.
 - [43] W. L. WAN AND T. F. CHAN, *A phase error analysis of multigrid methods for hyperbolic equations*, SIAM J. Sci. Comput., 25 (2003), pp. 857–880.
 - [44] D. L. WILLIAMSON, *The evolution of dynamical cores for global atmospheric models*, J. Meteorol. Soc. Jpn., 85B (2007), pp. 241–269.
 - [45] *XBraid: Parallel multigrid in time*. <http://llnl.gov/casc/xbraid>.
 - [46] I. YAVNEH, *Coarse-grid correction for nonelliptic and singular perturbation problems*, SIAM J. Sci. Comput., 19 (1998), pp. 1682–1699.
 - [47] I. YAVNEH, C. H. VENNER, AND A. BRANDT, *Fast multigrid solution of the advection problem with closed characteristics*, SIAM J. Sci. Comput., 19 (1998), pp. 111–125.

**SUPPLEMENTARY MATERIALS: FAST MULTIGRID
REDUCTION-IN-TIME FOR ADVECTION VIA MODIFIED
SEMI-LAGRANGIAN COARSE-GRID OPERATORS**

These supplementary materials are organized as follows. [Section SM1](#) presents supporting numerical evidence for [Lemma 3.3](#) and the claims made in [Remark 3.4](#). [Section SM2](#) describes an implementation detail relating to the use of GMRES to inexactly solve coarse-grid linear systems. [Section SM3](#) describes some further details about the variable-wave-speed test problems we consider in one dimension. [Section SM4](#) provides some additional parallel strong-scaling results for the one-dimensional problem. [Section SM5](#) describes some further details about the variable-wave-speed test problem we consider in two dimensions. Finally, [Section SM6](#) provides a strategy for estimating coarse-grid departure points in two spatial dimensions.

SM1. Supporting numerical evidence for [Lemma 3.3](#) and claims made in [Remark 3.4](#). In this section, we provide supporting numerical evidence for [Lemma 3.3](#) and the claims made in [Remark 3.4](#). To do so, we test numerically to what extent (if any) the relationships [\(3.14\)](#) and [\(3.16\)](#) hold when the wave-speed function varies in space. In particular, we consider the following two wave-speed functions

$$\begin{aligned} \text{(SM1)} \quad & \alpha(x, t) = \cos(2\pi t), \\ \text{(SM2)} \quad & \alpha(x, t) = \cos(2\pi t) \cos(2\pi x), \end{aligned}$$

and we integrate from time $t_n = 0$ to time $t_n + m\delta t$, taking $m = 4$. To mimic *exactly locating* departure points—as the idealized semi-Lagrangian schemes $\mathcal{S}_{p,\infty}^{(t_n+k\delta t,\delta t)}$ and $\mathcal{S}_{p,\infty}^{(t_n,m\delta t)}$ in [\(3.14\)](#) and [\(3.16\)](#) do—we use MATLAB’s `ode45` with very tight tolerances to integrate backwards along local characteristics with high accuracy.

Based on the results in [Lemma 3.3](#), numerically we measure the following quantity under mesh refinement in h :

$$\text{(SM3)} \quad \frac{\left\| \prod_{k=0}^{m-1} \mathcal{S}_{p,\infty}^{(t_n+k\delta t,\delta t)} \mathbf{u}(t_n) - C^{(t_n,m\delta t)} \mathcal{S}_{p,\infty}^{(t_n,m\delta t)} \mathbf{u}(t_n) \right\|_{\infty}}{\left\| \boldsymbol{\varphi}_{p+1}^{(t_n,m\delta t)} \right\|_{\infty}}.$$

Here, $C^{(t_n,m\delta t)}$ is either the identity matrix, or $C^{(t_n,m\delta t)} = \mathcal{B}_{p+1}^{(t_n,m\delta t)} := \left[I - \text{diag} \left(\boldsymbol{\varphi}_{p+1}^{(t_n,m\delta t)} \right) \mathcal{D}_{p+1} \right]^{-1}$, which can be thought of as a *correction* matrix that may map the rediscretized coarse-grid operator $\mathcal{S}_{p,\infty}^{(t_n,m\delta t)}$ closer to the ideal coarse-grid operator $\prod_{k=0}^{m-1} \mathcal{S}_{p,\infty}^{(t_n+k\delta t,\delta t)}$.

First, in [Figure SM1](#) we provide numerical verification of [Lemma 3.3](#). That is, we plot the quantity [\(SM3\)](#) for the spatially independent wave-speed function [\(SM1\)](#). If $C^{(t_n,m\delta t)} = I$, then from [\(3.14\)](#) we expect [\(SM3\)](#) to decay as $\mathcal{O}(h^{p+1})$, which is indeed what we observe, whether $\delta t = \mathcal{O}(h)$, or $\delta t = \mathcal{O}(1)$. Furthermore, if $C^{(t_n,m\delta t)} = \mathcal{B}_{p+1}^{(t_n,m\delta t)}$, then from [\(3.16\)](#) we expect [\(SM3\)](#) to decay as $\mathcal{O}(h^{p+2})$, which is indeed what we observe, whether $\delta t = \mathcal{O}(h)$, or $\delta t = \mathcal{O}(1)$.

Now, in [Figure SM2](#) we provide numerical evidence for the claims made in [Remark 3.4](#) regarding whether the spatially independent wave-speed estimates extend

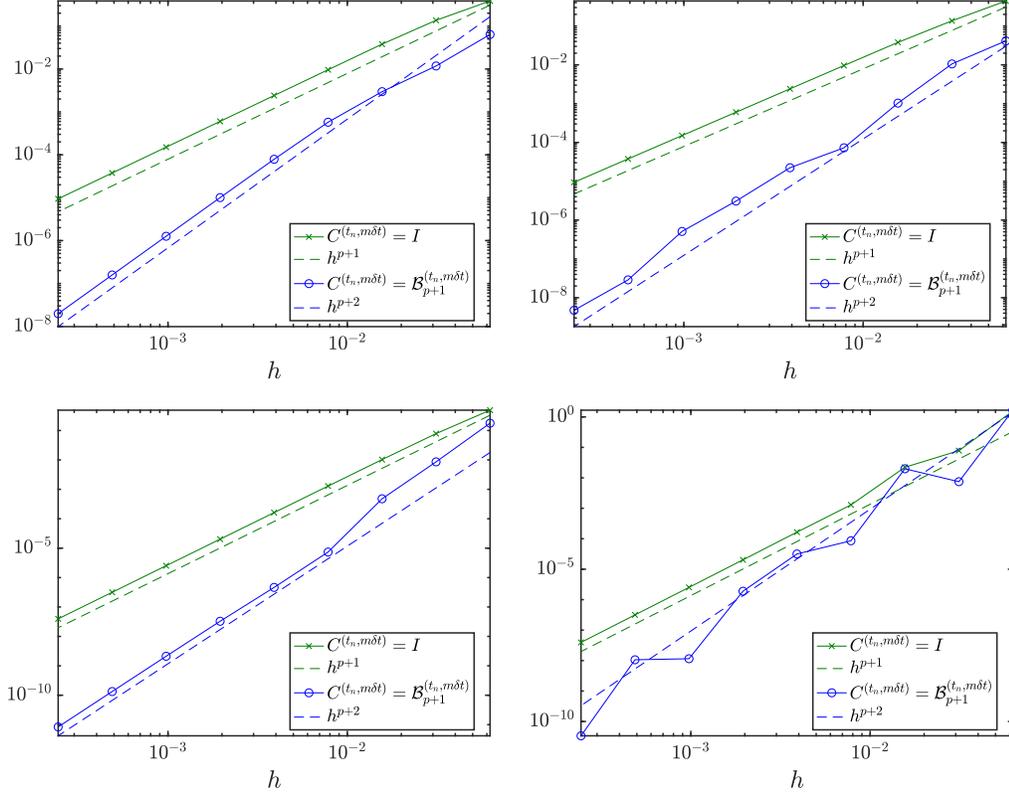


FIG. SM1. Numerical verification of Lemma 3.3. The quantity (SM3) with spatially independent wave-speed (SM1) is plotted under mesh refinement in h . Top: $p = 1$. Bottom $p = 2$. Left: $\delta t = 0.85h$. Right: $\delta t = 0.85$.

to spatially variable wave-speeds. That is, we test whether the relationships (3.14) and (3.16) hold when the wave-speed function is given by (SM2). Consider when $C^{(t_n, m\delta t)} = \mathcal{B}_{p+1}^{(t_n, m\delta t)}$, so as to verify whether the estimate (3.16) holds. For $p = 1$ (top row of Figure SM2), we see that (SM3) decays as $\mathcal{O}(h^{p+2})$ if $\delta t = \mathcal{O}(h)$ (top left panel), but decays as $\mathcal{O}(h^{p+1})$ if $\delta t = \mathcal{O}(1)$ (top right panel). Both of these results are consistent with estimate (3.16) holding up to terms of size $\mathcal{O}(h^{p+1}\delta t)$, as conjectured in Remark 3.4 when p is odd. However, for the $p = 2$ case (bottom row of Figure SM2), we see that when $C^{(t_n, m\delta t)} = \mathcal{B}_{p+1}^{(t_n, m\delta t)}$, the quantity (SM3) only decays as $\mathcal{O}(h^{p+1})$, even when $\delta t = \mathcal{O}(h)$. This is why, as stated in Remark 3.4, we do not believe that all of the estimates generalize to the spatially variable case when p is even.

SM2. Nonlinearity introduced by GMRES. Using GMRES to inexactly solve the coarse-grid linear systems introduces nonlinearity to the problem, since the approximations generated by GMRES depend nonlinearly on the right-hand sides of the linear systems. Applying GMRES inexactly within the full approximation scheme (FAS) framework is not straightforward (though possible), because it is important to ensure that the same GMRES polynomial is used for both the coarse-grid solve and the construction of the tau-correction term. To avoid this complication, we use a linear version of MGRIT, which we have implemented in XBraid (by default, XBraid uses the

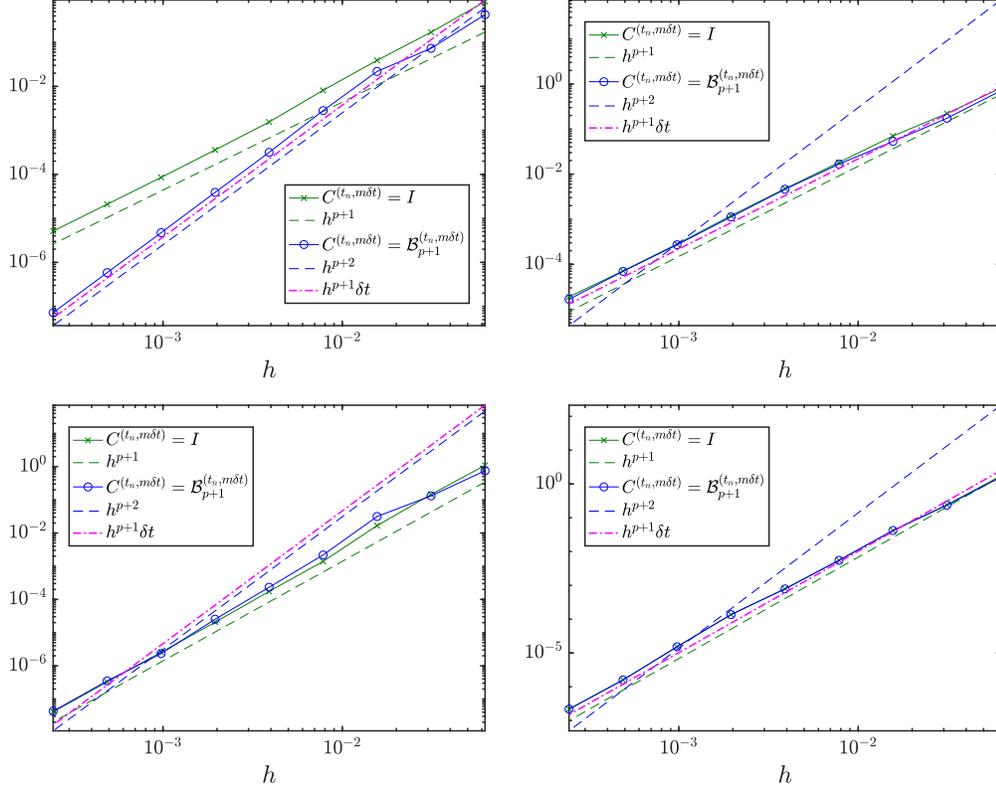


FIG. SM2. Supporting numerical evidence for Remark 3.4. The quantity (SM3) with spatially variable wave-speed (SM2) is plotted under mesh refinement in h . Top: $p = 1$. Bottom $p = 2$. Left: $\delta t = 0.85h$. Right: $\delta t = 0.85$.

FAS framework). This has the added benefit of reducing memory and computation overhead. Note that an FAS implementation of our algorithm would not produce identical results to the linear implementation used here, since the initial GMRES residuals (and hence the GMRES polynomials) would be different. The impact of this difference on convergence, if any, is a topic of future work.

SM3. One-dimensional numerical test problems. In this Section, plots of the wave-speed functions used in the numerical tests for the one-dimensional advection problem are shown. In addition, plots of the associated PDE solutions are shown. Figure SM3 presents the case in which the wave-speed function depends on time only. Figure SM4 presents the case in which the wave-speed function depends on both space and time.

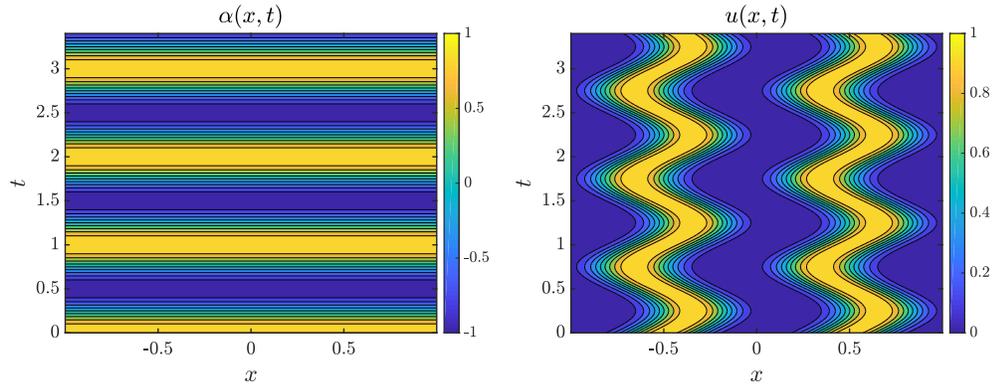


FIG. SM3. *Left: Wave-speed function $\alpha(x,t) = \cos(2\pi t)$, as given in (3.24). Right: The corresponding space-time solution of the one-dimensional advection problem (2.1). Note that the wave-speed function and PDE solution are shown on the truncated time domain $t \in [0, 3.4]$ rather than the full domain $t \in [0, 13.6]$ used in the numerical tests since they are easier to visualize on this shorter domain.*

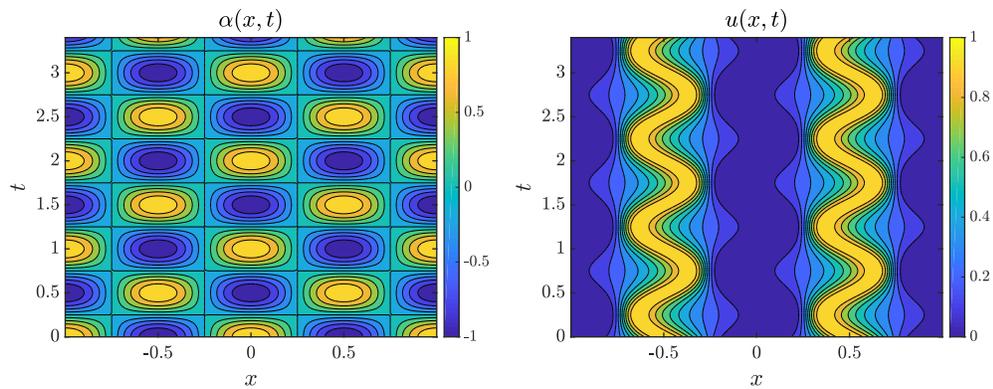


FIG. SM4. *Left: Wave-speed function $\alpha(x,t) = \cos(2\pi t) \cos(2\pi x)$, as given in (3.25). Right: The corresponding space-time solution of the one-dimensional advection problem (2.1). Note that the wave-speed function and PDE solution are shown on the truncated time domain $t \in [0, 3.4]$ rather than the full domain $t \in [0, 13.6]$ used in the numerical tests since they are easier to visualize on this shorter domain.*

SM4. One-dimensional strong-scaling tests solving to discretization accuracy. In the left panel of Figure SM5 we reproduce the strong-scaling plot from the left panel of Figure 5 for the one-dimensional advection problem. As throughout the rest of the paper, these tests iterate MGRIT until the ℓ^2 -norm of the space-time residual is reduced by at least 10 orders of magnitude from its initial value. This required 13, 19, and 24 MGRIT iterations for $(p, r) = ((1, 1), (3, 3), (5, 5))$, respectively. In the right panel of Figure SM5 we show strong-scaling results corresponding to iterating MGRIT until the solution at the final time point reaches discretization error accuracy in the discrete ℓ^2 -norm. The exact PDE solution needed to measure the discretization error is given in [27, App. B.4]. For $(p, r) = ((1, 1), (3, 3), (5, 5))$, the discretization error at the final time is approximately $0.75, 1.0 \times 10^{-5}$, and 8.6×10^{-10} , respectively, and requires 5, 12, and 25 MGRIT iterations to reach this. Since solving to discretization error accuracy requires fewer iterations than reducing the residual norm by 10 orders of magnitude for the 1st- and 3rd-order discretizations (i.e., reducing the residual norm by 10 order of magnitude *over solves* these problems), the speed-ups are slightly better.

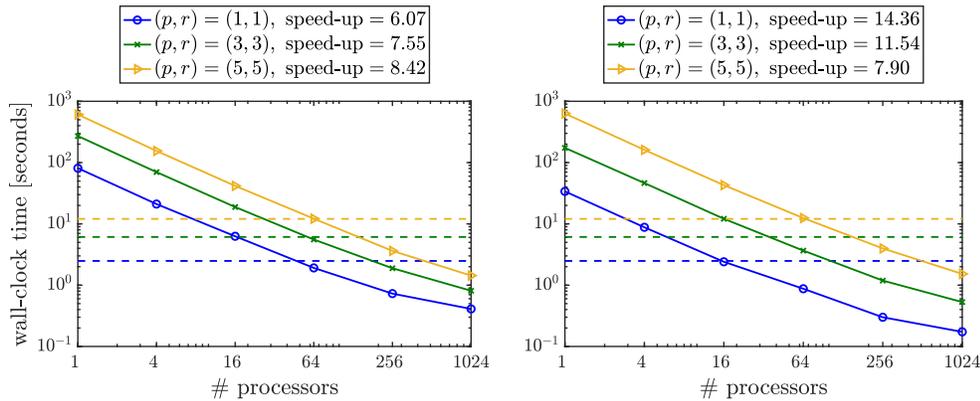


FIG. SM5. *Strong parallel scaling using time-only parallelism: Runtimes of MGRIT V-cycles with an aggressive coarsening factor of $m = 16$ on the first level followed by $m = 4$ on all coarser levels. The one-dimensional PDE (2.1) uses wave-speed (3.25), and the semi-Lagrangian discretizations of orders $(p, r) = ((1, 1), (3, 3), (5, 5))$ use a space-time mesh having $n_x \times n_t = 2^{12} \times 2^{14}$ points. Left: Fixed space-time residual stopping tolerance of 10^{-10} . Right: Residual stopping tolerance based on reaching discretization error at the final time point. Dashed lines represent runtimes of time-stepping on one processor. Speed-ups obtained using 1024 processors are listed in the legends.*

SM5. Two-dimensional numerical test problem. Plots of the velocity field associated with the variable-wave-speed function (4.9) are shown in Figure SM6 for several different times. In addition, snapshots of the associated PDE solution are shown for several different times in Figure SM7.

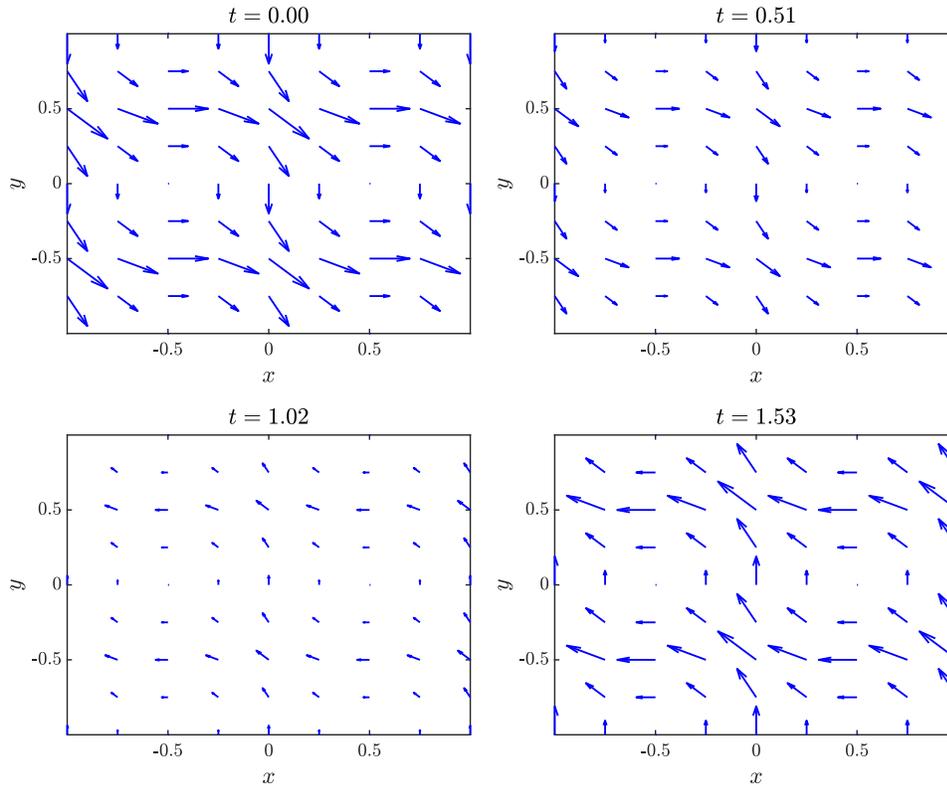


FIG. SM6. Plotted at four different times $t = (0, 0.51, 1.02, 1.53)$ is the velocity field $(\alpha(x, y, t), \beta(x, y, t))$ of the two-dimensional advection problem (4.1) for wave-speed given by (4.9). The four times $t = (0, 0.51, 1.02, 1.53)$ represent 0%, 15%, 30%, and 45% of one period of the time-periodic velocity field, which has a period of 3.4.

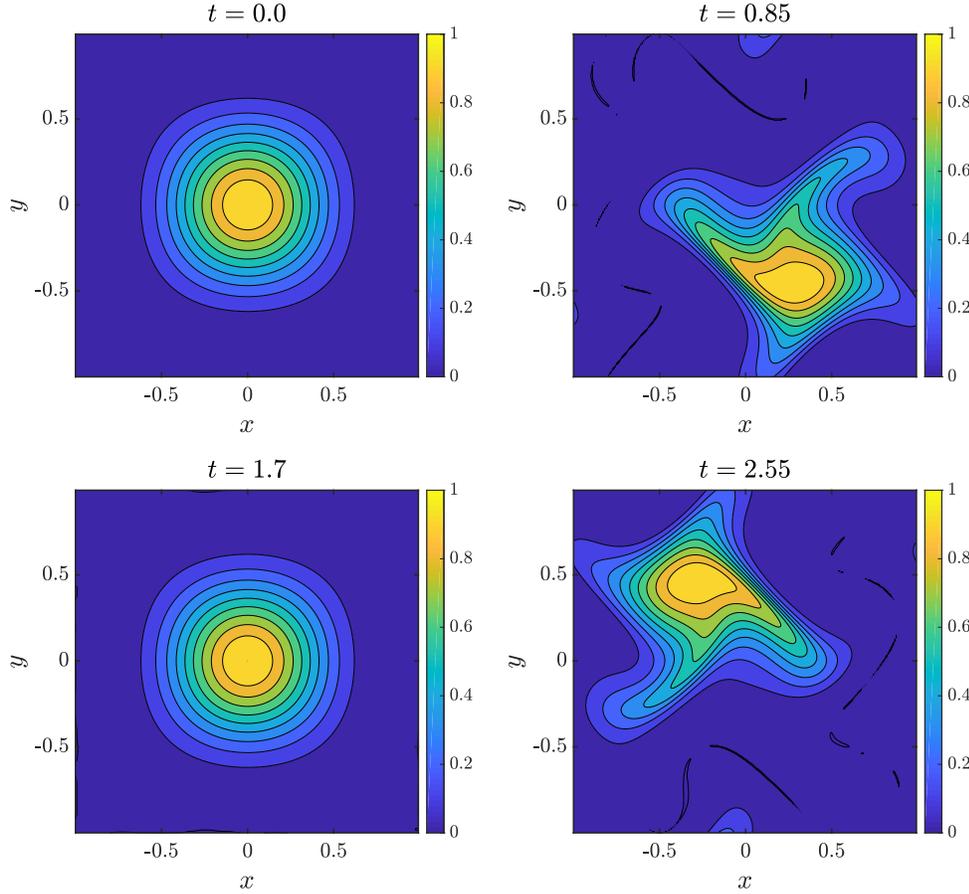


FIG. SM7. One period of the time-periodic solution of the two-dimensional advection problem (4.1) with wave-speed given by (4.9). Moving left to right, top to bottom, the solution is shown at the times $t = (0, 0.85, 1.70, 2.55, 3.4)$, representing 0%, 25%, 50%, and 75% of one period, respectively. Note: These snapshots are taken from a simulation that used a 5th-order discretization with $n_x^2 = 256^2$ points in space.

SM6. Estimating coarse-grid departure points in two dimensions via backtracking and linear interpolation. In this section, we propose a strategy for approximating coarse-grid departure points in two dimensions by a backtracking and linear interpolation procedure, generalizing the one-dimensional procedure proposed in Section 3.3.1. In Section SM6.1 we first present the bilinear interpolating polynomial that is used in this procedure, then Section SM6.2 presents the details of procedure.

SM6.1. A two-dimensional interpolating polynomial. In one spatial dimension, define the linear interpolating polynomial $p^{1d}(x; q_E, q_W)$ as that which interpolates the two values q_E and q_W associated with the two mesh points $x = x_W$, and $x = x_E$, respectively, with $x_E = x_W + h$. For example, q_E and q_W could represent the evaluation of some function $q(x)$ at $x = x_E$ and $x = x_W$. This interpolating

polynomial can be written as

$$(SM1) \quad p^{1d}(x; q_E, q_W) = \frac{q_E - q_W}{h}(x - x_E) + q_E, \quad x \in [x_W, x_E].$$

Using the one-dimensional polynomial (SM1) we now construct a two-dimensional interpolating polynomial. Let $p^{2d}(x, y; r_{NE}, r_{NW}, r_{SE}, r_{SW})$ denote the bilinear function that interpolates the four values $r_{NE}, r_{NW}, r_{SE}, r_{SW}$ that are associated with the four mesh points $(x, y) = ((x_E, y_N), (x_W, y_N), (x_E, y_S), (x_W, y_S))$, respectively, with $y_N = x_S + h$. In terms of (SM1), this function can be written as

$$(SM2) \quad \begin{aligned} & p^{2d}(x, y; r_{NE}, r_{NW}, r_{SE}, r_{SW}) \\ &= \frac{p^{1d}(x; r_{NE}, r_{NW}) - p^{1d}(x; r_{SE}, r_{SW})}{h}(y - y_N) + p^{1d}(x; r_{NE}, r_{NW}), \end{aligned}$$

with $(x, y) \in [x_W, x_E] \times [y_S, y_N]$.

SM6.2. Coarse-grid departure point estimation. To begin, it is helpful to recall the notation used in Section 4 to describe a local characteristic. Recall that the characteristic

$$(SM3) \quad (x(t), y(t)) = (\xi_{ij}^{(t_n, \delta t)}(t), \eta_{ij}^{(t_n, \delta t)}(t)),$$

is defined over the interval $t \in [t_n, t_n + \delta t]$. The *arrival* point of this characteristic is $(\xi_{ij}^{(t_n, \delta t)}(t_n + \delta t), \eta_{ij}^{(t_n, \delta t)}(t_n + \delta t)) = (x_i, y_j)$, and the departure point is its location at time $t = t_n$.

As in Section 3.3.1, for notational simplicity, suppose that we are working on the first coarse-grid time interval, $t \in [0, m\delta t]$, but note that the following strategy can be applied on any coarse-grid time interval $t \in [t_n, t_n + m\delta t]$. Consider the local coarse-grid characteristic

$$(SM4) \quad (x(t), y(t)) = (\xi_{ij}^{(0, m\delta t)}(t), \eta_{ij}^{(0, m\delta t)}(t))$$

that arrives at the mesh point $(x, y) = (x_i, y_j)$ at time $t = m\delta t$. Our objective is to estimate the departure point of this characteristic: $(\xi_{ij}^{(0, m\delta t)}(0), \eta_{ij}^{(0, m\delta t)}(0))$.

Generalizing our one-dimensional strategy from Section 3.3.1, we propose to estimate this departure point by approximately tracking the path of the characteristic (SM4) backwards from $t = m\delta t \rightarrow t = 0$. More specifically, we do so by using the local fine-grid characteristic directions that have already been computed on fine-grid subintervals $t \in [k\delta t, (k+1)\delta t], k = m-1, \dots, 0$ to guide its path in an interpolating manner. Let our approximation to the coarse-grid characteristic at time $t = k\delta t$ be denoted by

$$(SM5) \quad (c_{ij}^{(k)}, d_{ij}^{(k)}) \approx (\xi_{ij}^{(0, m\delta t)}(k\delta t), \eta_{ij}^{(0, m\delta t)}(k\delta t)).$$

The approximate departure point is therefore denoted by $(c_{ij}^{(0)}, d_{ij}^{(0)})$.

Since (SM4) arrives at $(x, y) = (x_i, y_j)$ at time $t = m\delta t$, over the last fine-grid subinterval $t \in [(m-1)\delta t, m\delta t]$ it is the same as the local fine-grid characteristic that also arrives at $(x, y) = (x_i, y_j)$ at time $t = m\delta t$. As such, we use the departure point of this fine-grid characteristic as a final-time condition to initialize our approximation:

$$(SM6) \quad (c_{ij}^{(m-1)}, d_{ij}^{(m-1)}) = (\xi_{ij}^{((m-1)\delta t, \delta t)}((m-1)\delta t), \eta_{ij}^{((m-1)\delta t, \delta t)}((m-1)\delta t)).$$

Given $(c_{ij}^{(k+1)}, d_{ij}^{(k+1)})$, for a single $k \in \{m-2, \dots, 0\}$, we now describe how to compute $(c_{ij}^{(k)}, d_{ij}^{(k)})$; that is, how to propagate the coarse-grid characteristic from $t = (k+1)\delta t \rightarrow t = k\delta t$. Define $(x, y) = ((x_E, y_N), (x_W, y_N), (x_E, y_S), (x_W, y_S))$ as the four mesh points that are immediately to the north-east, north-west, south-east, and south-west, respectively, of $(x, y) = (c_{ij}^{(k+1)}, d_{ij}^{(k+1)})$. See the schematic in [Figure SM8](#). Furthermore, denote the four local fine-grid characteristics that *arrive* at these four mesh points by

$$(SM7) \quad (x(t), y(t)) = (\xi_{NE}^{(k\delta t, \delta t)}(t), \eta_{NE}^{(k\delta t, \delta t)}(t)),$$

$$(SM8) \quad (x(t), y(t)) = (\xi_{NW}^{(k\delta t, \delta t)}(t), \eta_{NW}^{(k\delta t, \delta t)}(t)),$$

$$(SM9) \quad (x(t), y(t)) = (\xi_{SE}^{(k\delta t, \delta t)}(t), \eta_{SE}^{(k\delta t, \delta t)}(t)),$$

$$(SM10) \quad (x(t), y(t)) = (\xi_{SW}^{(k\delta t, \delta t)}(t), \eta_{SW}^{(k\delta t, \delta t)}(t)),$$

respectively. Then, to approximate the x -component of the coarse-grid characteristic at time $t = k\delta t$, we simply fit a bilinear function to how the fine-grid characteristics [\(SM7\)](#)–[\(SM10\)](#) map the x -component of their arrival point into the x -component of their departure point:

$$(SM11) \quad c_{ij}^{(k)} = p^{2d}(c_{ij}^{(k+1)}, d_{ij}^{(k+1)}; \xi_{NE}^{(k\delta t, \delta t)}(k\delta t), \xi_{NW}^{(k\delta t, \delta t)}(k\delta t), \xi_{SE}^{(k\delta t, \delta t)}(k\delta t), \xi_{SW}^{(k\delta t, \delta t)}(k\delta t)).$$

To estimate the y -component of the coarse-grid characteristic at time $t = k\delta t$, we

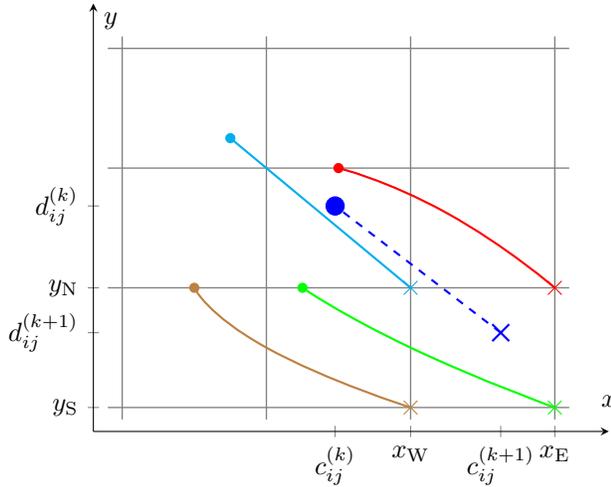


FIG. SM8. A bird's eye view of the four local fine-grid characteristics [\(SM7\)](#), [\(SM8\)](#), [\(SM9\)](#), [\(SM10\)](#) (the solid red, cyan, green, and brown lines, respectively). These characteristics arrive at time $t = (k+1)\delta t$ at the mesh points (small crosses) that are the closest neighbors of the point $(c_{ij}^{(k+1)}, d_{ij}^{(k+1)})$. The departure points of these local fine-grid characteristics at time $t = k\delta t$ are marked as small circles. The point $(c_{ij}^{(k+1)}, d_{ij}^{(k+1)})$ (large blue cross) is the approximate location of the coarse-grid characteristic [\(SM4\)](#) at time $t = (k+1)\delta t$. As described in [Section SM6.2](#), the unknown point $(c_{ij}^{(k)}, d_{ij}^{(k)})$ (large blue circle) is the approximate location of the coarse-grid characteristic [\(SM4\)](#) at time $t = k\delta t$, and is obtained by fitting a bilinear function to the paths of these neighboring fine-grid characteristics.

carry out the analogous procedure of fitting a bilinear function to how the fine-grid characteristics (SM7)–(SM10) map the y -component of their arrival point into the y -component of their departure point:

$$(SM12) \quad \begin{aligned} d_{ij}^{(k)} &= p^{2d}(c_{ij}^{(k+1)}, d_{ij}^{(k+1)}); \\ &\eta_{NE}^{(k\delta t, \delta t)}(k\delta t), \eta_{NW}^{(k\delta t, \delta t)}(k\delta t), \eta_{SE}^{(k\delta t, \delta t)}(k\delta t), \eta_{SW}^{(k\delta t, \delta t)}(k\delta t). \end{aligned}$$