

ADDITIVE POLYNOMIAL TIME INTEGRATORS, PART I:
FRAMEWORK AND FULLY IMPLICIT-EXPLICIT COLLOCATION
METHODS*TOMMASO BUVOLI[†] AND BEN S. SOUTHWORTH[‡]

Abstract. In this paper we generalize the polynomial time integration framework to additively partitioned initial value problems. The framework we present is general and enables the construction of many new families of additive integrators with arbitrary order-of-accuracy and varying degree of implicitness. In this first work, we focus on a new class of implicit-explicit polynomial block methods that are based on fully implicit Runge–Kutta methods with Radau nodes and that possess high stage order. We show that the new fully implicit-explicit (FIMEX) integrators have improved stability compared to existing IMEX Runge–Kutta methods, while also being more computationally efficient due to recent developments in preconditioning techniques for solving the associated systems of nonlinear equations. For PDEs on periodic domains where the implicit component is trivial to invert, we will show how parallelization of the right-hand side evaluations can be exploited to obtain significant speedup compared to existing serial IMEX Runge–Kutta methods. For parallel (in space) finite element discretizations, the new methods can achieve orders of magnitude better accuracy than existing IMEX Runge–Kutta methods and/or achieve a given accuracy several times faster in terms of computational runtime.

Key words. additive integrators, linearly implicit, implicit-explicit, fully implicit Runge–Kutta, general linear methods

MSC codes. 65L04, 65L05, 65L06

DOI. 10.1137/21M1446988

1. Introduction. Many problems in science and engineering can be modeled using high-dimensional systems of ordinary differential equations. These equations typically arise from the mathematical description of physical phenomena or from the spatial discretization of partial differential equations. Solving these systems amounts to integrating an initial value problem (IVP)

$$(1.1) \quad y'(t) = f(t, y(t)), \quad y(t_0) = y_0.$$

In practice, it is common to additively partition the right-hand side, $f(t, y)$, into m components,

$$(1.2) \quad f(t, y) = \sum_{k=1}^m f^{\{k\}}(t, y).$$

A simple example is a discretized advection-diffusion-reaction equation where each physical process is represented by a separate term.

*Submitted to the journal’s Methods and Algorithms for Scientific Computing section September 17, 2021; accepted for publication (in revised form) July 14, 2023; published electronically November 29, 2023.

<https://doi.org/10.1137/21M1446988>

Funding: The second author was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory as a Nicholas C. Metropolis Fellow and under project 20220174ER. The first author was funded in part by the NSF grant DMS-2012875.

[†]Department of Mathematics, Tulane University, New Orleans, LA 70118 USA (tbuvoli@tulane.edu).

[‡]Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545 USA (southworth@lanl.gov).

The solution of any additively partitioned system can be numerically approximated using an additive integrator that treats each component $f^{\{k\}}(t, y)$ differently [22, 33, 40]. This can be particularly efficient for solving multiscale, multiphysics problems where the optimal method differs across the components, or where it is prohibitively expensive to treat the full operator $f(t, y)$ implicitly. A canonical example is a linear advection-diffusion equation, where the diffusion places severe explicit timestep restrictions, but fully implicit solves for advection-diffusion discretizations are significantly more challenging than for pure diffusion. An additive integrator can treat the diffusion implicitly and the advection explicitly, addressing each of these problems. If the right-hand side consists of two components and an additive integrator treats the first implicitly and the second explicitly, then the integrator is frequently called an implicit-explicit (IMEX) method.

Two closely related classes of integrators are linearly implicit methods and W-methods that, respectively, utilize an exact or approximate local Jacobian of $f(t, y)$ at each timestep [29, Chapter IV, Section 7] [15, 1, 27]. Given the solution at the n th timestep $y_n = y(t_n)$, we can rewrite the system (1.1) as an additively partitioned system with $m = 2$,

$$(1.3) \quad f^{\{1\}}(t, y) = J_n y, \quad \text{and} \quad f^{\{2\}}(t, y) = f(t, y) - J_n y,$$

where J_n approximates or is equal to the local Jacobian $\frac{\partial f}{\partial y}(t_n, y_n)$. Any additive integrator that treats $f^{\{1\}}(t, y)$ implicitly and $f^{\{2\}}(t, y)$ explicitly reduces to a linearly implicit method.

In the past three decades, the construction of additive integrators has been an active area of research that has produced a range of methods, including linear multistep methods (LMMs) [6, 23, 46], Runge–Kutta (RK) methods [5, 33, 37, 40, 31, 34], and general linear methods (GLMs) [17, 48, 43], including those based on extrapolation [21, 16]. Each method class has certain benefits and drawbacks. Additive LMMs have a low computational cost per timestep, but high-order methods experience instabilities on equations with limited diffusion. Additive diagonally implicit RK methods possess good stability and allow for simplified adaptive timestepping, but they are known to suffer from order-reduction on stiff equations [7]. Moreover, high-order RK method derivations that rely on nonlinear order conditions grow increasingly difficult to construct, and alternative approaches must be considered [21, 37]. GLMs with good stability and no order-reduction exist, but to avoid increasing the number of order conditions further, one must typically consider simplified method formulations.

In this paper we will generalize the recently introduced polynomial time integration framework [9, 14, 10, 13] to include additively partitioned differential equations. We then demonstrate its utility by introducing a new family of IMEX integrators where the implicit integrator is a fully implicit collocation method. Selecting a fully implicit integrator may seem peculiar since these methods are often considered too slow to be competitive. However, recent developments in block linear and nonlinear solvers make their use in numerical PDEs quite tractable [18, 38, 26, 32, 39, 45, 44], even outperforming diagonally implicit RK methods in many cases [45, 44]. Despite these developments, it is not possible to derive IMEX-RK methods based on the fully implicit RK methods since the explicit stages would be nonlinearly coupled to the fully implicit stages. More generally, constructing high-order IMEX-RK schemes is often nontrivial, as mentioned in [34]. In this work, we will show how the additive polynomial framework provides a natural way to develop fully implicit-explicit (FIMEX) integrators, which are high-order accurate, and allow us to leverage developments in fully implicit solvers in the context of additive integration.

More generally, there are two main advantages of the additive polynomial framework that will be explored in this paper. First, the framework simplifies the construction of high-order additive GLMs that do not suffer from order-reduction on stiff equations. Specifically, the polynomial framework makes extensive use of interpolating polynomials that trivially satisfy nonlinear order conditions and ensure high stage order. Furthermore, method construction can be done using geometric arguments that are similar to those used to derive spatial finite difference stencils. Second, the framework can be used to derive efficient methods for solving equations that are either naturally split into multiple terms or where the right-hand side has been rewritten using an exact or approximate Jacobian. This allows us to simultaneously introduce a range of new high-order additive and semi-implicit integrators.

This paper is organized as follows. In sections 2 and 3 we, respectively, provide short introductions to additive integrators and the polynomial framework. In section 4 we generalize the polynomial framework for additively partitioned differential equations, and in section 5 we develop new classes of FIMEX methods and study their stability. Last, section 6 demonstrates the improved accuracy and efficiency of the new integrators for solving partial differential equations.

2. Additive integrators. Additive integrators are a class of methods for solving the partitioned initial value problem (1.1), (1.2). In this section, we give a short introduction to additive integrators for equations with two partitions ($m = 2$), where

$$(2.1) \quad y' = f^{\{1\}}(t, y) + f^{\{2\}}(t, y), \quad y(t_0) = y_0.$$

We assume that the term $f^{\{1\}}(t, y)$ is stiff (i.e., a small stepsize is required for any explicit method when solving $y' = f^{\{1\}}(t, y)$), while the term $f^{\{2\}}(t, y)$ is nonstiff.

In such a scenario it is desirable to consider integrators that only treat $f^{\{1\}}(t, y)$ implicitly. One approach for deriving additive methods is to integrate (2.1) and then approximate the resulting integrals for each term separately,

$$(2.2) \quad y(t_{n+1}) = y(t_n) + \underbrace{\int_{t_n}^{t_{n+1}} f^{\{1\}}(t, y(t)) dt}_{\text{treat implicitly}} + \underbrace{\int_{t_n}^{t_{n+1}} f^{\{2\}}(t, y(t)) dt}_{\text{treat explicitly}}.$$

One of the simplest additive integrators can be derived by taking an implicit one-sided approximation for $f^{\{1\}}(t, y)$, and an explicit one-sided approximation for $f^{\{2\}}(t, y)$. This produces the IMEX-Euler method

$$(2.3) \quad y_{n+1} = y_n + h f_{n+1}^{\{1\}} + h f_n^{\{2\}},$$

where the stepsize $h = t_{n+1} - t_n$. Higher-order IMEX-LMM methods [6, 23] use higher-order polynomial approximations constructed from previous solution values, while the output of a higher-order IMEX-RK method [5, 33, 40, 34] is a linear combination of newly computed stage values. IMEX-GLMs [17, 48, 43] combine both ideas by using previous solution values and new stages.

When using any IMEX method there are a range of choices for $f^{\{1\}}$ and $f^{\{2\}}$ that affect the computational cost and stability of the integrator. For a semilinear system $y' = Ly + N(t, y)$, with a nonstiff nonlinearity, it is natural to let

$$(2.4) \quad f^{\{1\}}(t, y) = Ly, \quad f^{\{2\}}(t, y) = N(t, y).$$

For a more general nonlinear system $y' = A(t, y) + B(t, y)$ that naturally splits into two components, $A(t, y)$ and $B(t, y)$, several choices for the numerical splitting include the following:

1. Fully implicit in $A(t, y)$:

$$(2.5) \quad f^{\{1\}}(t, y) = A(t, y), \quad f^{\{2\}}(t, y) = B(t, y).$$

2. Linearly implicit in $A(t, y)$:

$$(2.6) \quad f^{\{1\}}(t, y) = \frac{\partial A}{\partial y}(t_n, y_n)y, \quad f^{\{2\}}(t, y) = B(t, y) + A(t, y) - \frac{\partial A}{\partial y}(t_n, y_n)y.$$

3. Linearly implicit in $A(t, y)$ and $B(t, y)$:

$$(2.7) \quad \begin{aligned} f^{\{1\}}(t, y) &= \left[\frac{\partial A}{\partial y}(t_n, y_n) + \frac{\partial B}{\partial y}(t_n, y_n) \right] y, \\ f^{\{2\}}(t, y) &= B(t, y) + A(t, y) - \left[\frac{\partial A}{\partial y}(t_n, y_n) + \frac{\partial B}{\partial y}(t_n, y_n) \right] y. \end{aligned}$$

4. Linearly implicit in J_n (J_n approximates the full or partial Jacobian at $t = t_n$):

$$(2.8) \quad f^{\{1\}}(t, y) = J_n y, \quad f^{\{2\}}(t, y) = B(t, y) + A(t, y) - J_n y.$$

For additional clarity we write the formulas for the IMEX-Euler method (2.3) using three of the proposed partitionings:

$$\text{partitioning (2.4)} \quad y_{n+1} = (I - hL)^{-1}(y_n + hN(t_n, y_n)),$$

$$\text{partitioning (2.5)} \quad y_{n+1} = y_n + hA(t_{n+1}, y_{n+1}) + hB(t_n, y_n),$$

$$\text{partitioning (2.8)} \quad y_{n+1} = (I - hJ_n)^{-1}(y_n + hA(t_n, y_n) + hB(t_n, y_n)).$$

Selecting a fully implicit partitioning generally leads to improved stability but requires a nonlinear solve at each timestep. Conversely, a linearly implicit choice only requires a linear solve at each step, but the method may have inferior stability properties, especially if J_n does not closely approximate the local Jacobian.

In the sections that follow we will first review the polynomial time integration framework and then generalize it so that we can construct new high-order additive integrators for (2.1) using any of the proposed partitionings (2.4)–(2.8).

3. Polynomial time integrators. The polynomial time integration framework [9, 14] is based on continuous polynomials in time that are constructed by fitting through solution or derivative values. The values may be known (i.e., input values) or unknown (i.e., future stage values or outputs), with the latter leading to implicit equations and ultimately implicit methods. Within the family of classical time integration methods, we can interpret polynomial time integrators as parametrized GLMs whose inputs and outputs approximate the solution at a set of scaled nodes $\{z_j\}$. The nodes z_j may be positive, negative, or zero, and the scaling factor is a positive number called the *node radius* r . Note that the node radius introduces an additional degree of freedom that complements the stepsize h .

The input and output values of a polynomial integrator, along with the associated derivatives, are represented using the notation

$$\begin{aligned} \text{input (solutions):} \quad & y_j^{[n]} \approx y(t_n + rz_j), \\ \text{output (solutions):} \quad & y_j^{[n+1]} \approx y(t_n + rz_j + h), \\ \text{input derivatives:} \quad & f_j^{[n]} := f(t_n + rz_j, y_j^{[n]}) \approx y'(t_n + rz_j), \\ \text{output derivatives:} \quad & f_j^{[n+1]} := f(t_n + rz_j + h, y_j^{[n+1]}) \approx y'(t_n + rz_j + h). \end{aligned}$$

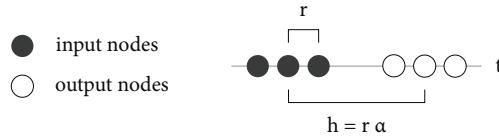


FIG. 1. A diagram showing the input and output nodes for a polynomial integrator with three equispaced input and output nodes (e.g., $z_j = \{-1, 0, 1\}$). The node set is scaled by the node radius and the distance between the inputs and outputs is proportional to the extrapolation factor α .

Since the polynomials that make up polynomial methods are expressed in local coordinates (3.2), it is convenient to also parametrize the stepsize h in terms of the node radius. We therefore let

$$(3.1) \quad h = r\alpha,$$

where the constant α is called the *extrapolation factor*. In Figure 1 we show a visualization of the parameters r , h , and α for a method with three real-valued, equispaced nodes $\{z_j\}$.

In this work we look at methods with real-valued quadrature nodes $\{z_j\}$ that are scaled relative to the interval $[-1, 1]$. This choice is inspired by the fact that nearly all of the theoretical analysis relating to orthogonal polynomials and interpolation is conducted on this interval. However, when implementing a polynomial integrator it is typically simpler to translate the nodes to the interval $[0, 2]$ such that the leftmost node at $\tau = 0$ is located at $t = t_0$, where the initial condition for (1.1) is provided. Translating the nodes only affects temporal locations of the inputs, as determined by (3.2), and has no effect on the method coefficients.

The name polynomial integrator originates from the fact that all the method coefficients are derived using interpolating polynomials, known as *ODE polynomials*. All ODE polynomials approximate the Taylor series of the solution in local coordinates τ , where

$$(3.2) \quad t(\tau) = r\tau + t_n.$$

The general form for an ODE polynomial of degree g with expansion point b is

$$(3.3) \quad p(\tau; b) = \sum_{j=0}^g \frac{a_j(b)(\tau - b)^j}{j!},$$

where the constants $\{a_j(b)\}$ are called *approximate derivatives* since they approximate the derivatives of the solution in local coordinates such that $a_j(b) \approx \left. \frac{d^j}{d\tau^j} y(t(\tau)) \right|_{\tau=b} = r^j y^{(j)}(t(b))$ (the factor of r originates from the transformation into local coordinates (3.2)). Each approximate derivative $a_j(b)$ is computed by differentiating interpolating polynomials that are constructed using any subset of the method's inputs, outputs, and stages and the corresponding derivatives. The order of the polynomial method is directly related to the degree of its ODE polynomials and the order of the interpolating polynomials that are used to construct the associated approximate derivatives [9, sect. 3.6]. As a consequence, the order of accuracy is always bounded below by $\min(g_{\min}, \delta - 1)$, where g_{\min} is the minimum degree of the ODE polynomials and δ is the minimum degree of all the polynomials that determine the approximate derivatives.

A general formulation for the approximate derivatives $a_j(b)$ is described in [9, 14]; here we will only describe two important subfamilies:

1. *Adams ODE polynomials* are constructed using two Lagrange interpolating polynomials $L_y(\tau)$ and $L_f(\tau)$ that, respectively, satisfy solution values or derivative values at the method's input, output, or stage nodes. Specifically, $L_y(\tau) \approx y(t(\tau))$ interpolates at least one solution value, and $L_f(\tau) \approx ry'(t(\tau))$ is a polynomial of degree $g - 1$ that interpolates g derivative values. The approximate derivatives are then

$$(3.4) \quad a_0(b) = L_y(b) \quad \text{and} \quad a_j(b) = \frac{d^{j-1}L_f}{d\tau^{j-1}}(b), \quad j = 1, \dots, g.$$

By substituting (3.4) into (3.3) and noting that $p'(\tau; b) = L_f(\tau)$, we can express an Adams ODE polynomial in the equivalent integral form

$$(3.5) \quad p(\tau; b) = L_y(b) + \int_b^\tau L_f(\xi) d\xi.$$

We now see that an Adams ODE polynomial approximates the integral equation of an initial value problem where the expansion point b is the location of the initial condition in local coordinates.

2. *Backward difference formula (BDF) ODE polynomials* are constructed using a polynomial $H_y(\tau) \approx y(t(\tau))$ of degree g that interpolates g solution values and whose derivative $H'_y(\tau)$ interpolates a single derivative value. The approximate derivatives are given by

$$(3.6) \quad a_j(b) = \frac{d^j H_y}{d\tau^j}(b) \quad \Rightarrow \quad p(\tau; b) = H_y(\tau) \quad \forall b.$$

Note that BDF polynomials do not depend on the expansion point b .

When presenting polynomial methods, it is convenient to introduce a set containing all the data values that can be used to construct the interpolating polynomials that determine the approximate derivatives $a_j(b)$ in the ODE polynomial (3.3). This set is called the *ODE dataset*, and in the case of classical polynomial methods, it simply contains the method's inputs, outputs, and stages and their derivatives, along with the corresponding temporal nodes. An ODE dataset of size w is denoted as

$$(3.7) \quad D(r, t_n) = \{(\tau_j, y_j, r f_j)\}_{j=1}^w, \quad \text{where} \quad y_j \approx y(t(\tau_j)), \quad f_j = f(t(\tau_j), y_j).$$

Using all the previous definitions, the formula for any polynomial method with s stages and q outputs can be written compactly as

$$(3.8) \quad \begin{aligned} Y_i &= p_j(c_j(\alpha); b_j(\alpha)), & j &= 1, \dots, s, \\ y_j^{[n+1]} &= p_{j+s}(z_j + \alpha; b_{j+s}(\alpha)), & j &= 1, \dots, q, \end{aligned}$$

where Y_i denote stage values, $c_j(\alpha)$ are stage nodes in local coordinates, and $p_j(\tau; b)$ are ODE polynomials constructed from an ODE dataset of size $w = 2q + s$ that contains the method's inputs, stage values, and outputs. The associated temporal nodes of the method's ODE datasets (3.7) are, respectively,

$$(3.9) \quad \tau_j = \begin{cases} z_j, & 1 \leq j \leq q, \\ c_{j-q}(\alpha), & q + 1 \leq j \leq q + s, \\ z_{j-q-s} + \alpha, & q + s < j \leq 2q + s. \end{cases}$$

We remark that our definition of stages differs from the standard convention used for GLMs and RK methods. Specifically, if one recasts the method (3.8) as a GLM, then the outputs $y_j^{[n+1]}$ also double as additional stage values; in other words, (3.8) will be a GLM with $s + q$ stages. By avoiding the standard convention we allow for a more compact method definition and avoid method coefficient duplication.

Using the general formulation (3.8), it is possible to derive many different families of polynomial integrators. One example is *polynomial block methods* (PBMs) from [14], which are characterized by $s = 0$ and will be the primary focus of this paper due to their simpler structure. Additional examples include well-known time integrators like BDFs, Adams–Moulton methods, and collocation methods, which can all be expressed as PBMs with a fixed α . In the following subsection we show how to write fully implicit collocation methods in the polynomial framework. The resulting formulation will be used again in section 5 to construct new additive polynomial integrators.

3.1. Collocation methods and Radau IIA. Collocation methods [28, Chapter II, Section 7] are time integrators based on polynomial quadrature that can also be expressed as fully implicit RK methods. Well-known examples include the A-stable Gauss methods [8, sect. 342] and the L-stable Radau IIA methods [30] that, respectively, achieve orders of 2σ and $2\sigma - 1$, where σ is the number of stages. An additional benefit of these methods is that they both satisfy B-stability [29, 47].

Suppose that we seek an approximate solution of (1.1) at $t = t_n + h$ given an initial condition $y_n = y(t_n)$. To derive a collocation method, we can approximate the solution $y(t)$ using a polynomial $p_y(t)$ that satisfies the initial condition at $t = t_n$ and the differential equation at a set of m collocation points $t_{n,j}$ such that $p'_y(t_{n,j}) = f(t_{n,j}, p_y(t_{n,j}))$. These constraints lead to the fully implicit nonlinear system

$$(3.10) \quad p_y(t_{n,j}) = y_n + \int_{t_n}^{t_{n,j}} \underbrace{\sum_{j=1}^{\sigma} \ell_j(t) f(t_{n,j}, p_y(t_{n,j})) dt}_{p_f(t)}, \quad j = 1 \dots, m,$$

where $\ell_j(t) = \prod_{k \neq j} (t - t_{n,k}) / (t_{n,j} - t_{n,k})$ is the j th Lagrange basis polynomial and $p_f(t)$ is a Lagrange interpolating polynomial for the solution derivative $y'(t)$. If we compare the right-hand side of (3.10) to the Adams ODE polynomial integral formulation (3.5), we see that $p_y(t(\tau))$ is equivalent to an Adams ODE polynomial $p(\tau; b)$ with $b = \tau(0) = t_n$, $L_y(\tau) = y_n$, and $L_f(\tau) = r p_f(t(\tau))$. Therefore, by appropriately defining an ODE dataset and node set $\{z_j\}$, we can express any collocation method (3.10) as a one-step polynomial integrator (i.e., (3.8) with $q = 1$, $s = \sigma$) whose output is computed using a single Adams ODE polynomial.

However, for the purposes of this paper we will instead rewrite a collocation method as a multivalued PBM (i.e., (3.8) with $q > 1$ and $s = 0$) that advances the full “block” solution at the set of collocation points forward in time by h . To simplify the derivation we focus solely on a Radau IIA method with $\sigma - 1$ stages and show how this integrator can be reexpressed as a PBM with fixed α and $q = \sigma$. Though such a formulation may seem unnecessarily verbose, we will use the multivalued inputs (i.e., solutions at collocation points from the previous timestep) in section 5 to construct a new class of additive IMEX integrators.

We start by selecting the PBM nodes

$$(3.11) \quad \{z_j\}_{j=1}^q = \{-1, 2x_1 - 1, \dots, 2x_{q-1} - 1\},$$

where x_j is the j th zero of the polynomial $\frac{d^{q-2}}{dx^{q-2}}(x^{q-2}(x-1)^{q-1})$. Our first quadrature node is $z_1 = -1$ and the remaining nodes are the $q - 1$ Radau nodes scaled on the interval $x \in [-1, 1]$. In Figure 2(a), we illustrate the input and output nodes for a method with $q = 3$.

The outputs of the PBM will all be computed using a single Adams ODE polynomial (3.5). Since we want a method based on implicit Radau quadrature, we select

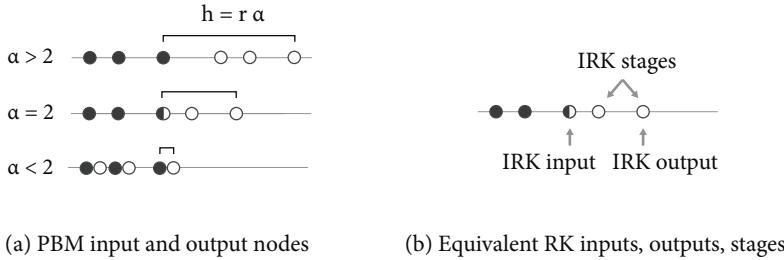


FIG. 2. Illustration of input nodes \bullet and output nodes \circ for the PBM with nodes (3.11) and $q = 3$, such that $\{z_j\} = \{-1, -1/3, 1\}$. The grey lines show the time axis which flows to the right. (a) An illustration of the temporal nodes for the input and outputs for three different values of α . When $\alpha = 2$ the last input node overlaps with the first output node. (b) A illustration relating the inputs and outputs of the PBM (3.12) with $b = 1$ and $\alpha = 2$ to the input, output, and stage values of an IRK method.

$L_f(\tau)$ to be the interpolating polynomial that matches all the output derivatives at the scaled Radau nodes ($L_f(z_j + \alpha) = f_j^{[n+1]}$ for $j = 2, \dots, q$). To maximize the accuracy of the integration constant $L_y(b)$ we choose $L_y(\tau)$ to be the interpolating polynomial that matches all the input values ($L_y(z_j) = y_j^{[n]}$ for $j = 1, \dots, q$). If we temporarily leave the expansion point b free, then we obtain the PBM

$$(3.12) \quad y_j^{[n+1]} = p(z_j + \alpha; b) = L_y(b) + \int_b^{z_j + \alpha} L_f(s) ds, \quad j = 1, \dots, q.$$

To obtain a one-step Radau IIA collocation we do the following:

1. *Select $\alpha = 2$.* The last input node now overlaps with the first output node; see Figure 2(a). This allows us to use the last input value as an accurate integration constant for Radau quadrature.
2. *Select $b = 1$.* This choice sets the last input $y_q^{[n]}$ as the integration constant since $L_y(1) = y_q^{[n]}$. Furthermore, $y_1^{[n+1]} = p(z_1 + 2; 1) = y_q^{[n]}$, and therefore the first output is equal to the last input.

The resulting PBM is equivalent to a $(2q - 3)$ th-order, one-step Radau IIA collocation method with stepsize $h = 2r$. Note that for general b and α , the PBM (3.12) is a $(q - 1)$ th-order, fully implicit multivalued integrator.¹

The inputs and outputs of the PBM with $\alpha = 2$ and $b = 1$ are related to the RK inputs and stages according to the following table, which is visualized in Figure 2(b).

RK step index ν	RK input y_ν	RK output $y_{\nu+1}$	RK stages
$n - 1$	$y_{n-1} = y_1^{[n]}$	$y_n = y_q^{[n]}$	$Y_{j-1} = y_j^{[n]}, j = 2, \dots, q$
n	$y_n = y_q^{[n]}$	$y_{n+1} = y_q^{[n+1]}$	$Y_{j-1} = y_j^{[n+1]}, j = 2, \dots, q$

For clarity we write the PBM for $q = 3$. The nodes are $\{z_j\} = \{-1, -1/3, 1\}$, and the polynomials are

$$L_y(\tau) = \frac{(t-1)(3t+1)}{4} y_1^{[n]} - \frac{9(t^2-1)}{8} y_2^{[n]} + \frac{(t+1)(3t+1)}{8} y_3^{[n]},$$

$$L_f(\tau) = -\frac{3}{4} (\tau - \alpha - 1) r f_2^{[n+1]} + \frac{3}{4} (\tau - \alpha + \frac{1}{3}) r f_3^{[n+1]}.$$

¹Regardless of the fact that we have Radau nodes, selecting general b will not produce a method with $2\sigma - 3$ accuracy. This is due to the fact that our initial condition will only be order $q - 1$ accurate, and the lower integration bound is different than the one used for Radau quadrature.

Selecting $b = 1$ and $\alpha = 2$, and substituting $L_y(\tau)$ and $L_f(\tau)$ into (3.12) yields

$$(3.13) \quad \begin{aligned} y_1^{[n+1]} &= y_3^{[n]}, \\ y_2^{[n+1]} &= y_3^{[n]} + \frac{5}{6}rf_2^{[n+1]} - \frac{1}{6}rf_3^{[n+1]}, \\ y_3^{[n+1]} &= y_3^{[n]} + \frac{3}{2}rf_2^{[n+1]} + \frac{1}{2}rf_3^{[n+1]}, \end{aligned}$$

which reduces to a Radau IIA method with stepsize $h = 2r$ (replacing r with $h/2$ produces the well-known IRK coefficients). Also note that the first two inputs are not used to compute the output, and therefore (3.13) is equivalent to a one-step method.

3.2. Iterators. Before introducing additive integrators, we require one additional concept from the polynomial framework, namely, the idea of an iterator. A polynomial integrator only advances the solution if the extrapolation factor α is greater than zero (see Figure 1). If we let $\alpha = 0$, then we obtain a special method known as an iterator. Iterators recompute the solution at the current timestep and certain method constructions share many similarities with predictor corrector block methods [41] and spectral deferred correction iterations [24, 19]. In [13] iterators were used in the context of exponential integration to compute initial conditions and create composite methods with improved stability. In this work, we will show how these same ideas can be applied to additive polynomial integrators.

4. Additive polynomial integrators. In this section we introduce the additive polynomial time integration framework for solving the partitioned system

$$(4.1) \quad y'(t) = f(t, y(t)) = \sum_{k=1}^m f^{\{k\}}(t, y(t)).$$

In subsection 4.1 we extend the ODE dataset (3.7) and the ODE polynomial (3.3) from the unpartitioned system (1.1) to the partitioned system (4.1). Then, in subsection 4.2 we introduce the class of additive PBMs that will provide a starting point for deriving new integrators in section 5.

4.1. Partitioned ODE datasets and ODE polynomials. A partitioned ODE dataset contains all the data values that can be used to construct partitioned ODE polynomials. We can trivially extend the ODE dataset for partitioned equations by replacing the full right-hand side $f(t, y)$ with all the derivative components $f^{\{j\}}(t, y)$.

DEFINITION 4.1 (partitioned ODE dataset). *A partitioned ODE dataset $D(r, s)$ of size w is an ordered set of tuples of the form*

$$(4.2) \quad D(r, s) = \left\{ \left(\tau_j, y_j, rf_j^{\{1\}}, \dots, rf_j^{\{m\}} \right) \right\}_{j=1}^w,$$

where $t(\tau) = r\tau + s$, $y_j \approx y(t(\tau_j))$, and $f_j^{\{k\}} = f^{\{k\}}(t(\tau_j), y_j)$.

Before generalizing the ODE polynomial it is convenient to first introduce the notion of a *total derivative*.

DEFINITION 4.2 (total derivative). *A total derivative F approximates the local ODE solution derivative $ry'(\tau)$ at a point $\tau = \hat{\tau}$, such that $F \approx ry'(t(\hat{\tau}))$, and is the sum of m derivative components, that is,*

$$(4.3) \quad F = rf_{j_1}^{\{1\}} + rf_{j_2}^{\{2\}} + \dots + rf_{j_m}^{\{m\}}, \quad \text{where } \hat{\tau} = \tau_{j_1} = \tau_{j_2} = \dots = \tau_{j_m},$$

for indices $j_k \in \{1, \dots, w\}$ and $k = 1, \dots, m$.

In other words, a total derivative is the sum of m component derivatives from an ODE dataset where the k th component derivative approximates $rf^{\{k\}}(t(\hat{\tau}), y(t(\hat{\tau})))$.

To generalize the ODE polynomial for partitioned systems, we can reuse (3.3) and simply modify the rules for computing approximate derivatives.

DEFINITION 4.3 (partitioned ODE polynomial). *A partitioned ODE polynomial of degree g with expansion point b can be expressed as*

$$(4.4) \quad p(\tau; b) = \sum_{j=0}^g \frac{a_j(b)(\tau - b)^j}{j!},$$

where each approximate derivative $a_j(b)$ is computed using values from a partitioned ODE dataset $D(r, s)$ in one of the following ways:

1. By differentiating a polynomial $h_j(\tau)$ that approximates $y(t(\tau))$; $h_j(\tau)$ must be a polynomial of least degree that interpolates at least one solution value in $D(r, s)$ and whose derivative $h'_j(\tau)$ interpolates any number of total derivatives from Definition 4.2. The j th approximate derivative of the ODE polynomial is then

$$a_j(b) = \left. \frac{d^j}{d\tau^j} h_j(\tau) \right|_{\tau=b}.$$

2. By differentiating a polynomial $l_j(\tau)$ that approximates $ry'(t(\tau))$; $l_j(\tau)$ is formed by summing m Lagrange interpolating polynomials so that

$$l_j(\tau) = \sum_{k=1}^m l_j^{\{k\}}(\tau),$$

where $l_j^{\{k\}}(\tau) \approx f^{\{k\}}(t(\tau), y(t(\tau)))$. Each of the interpolating polynomials $l_j^{\{k\}}$ must satisfy at least one of the derivative component values; specifically, for all k there exists at least one index $\nu \in \{1, \dots, w\}$, such that $l_j^{\{k\}}(\tau_\nu) = rf_\nu^{\{k\}}$. The j th approximate derivative is then

$$a_j(b) = \left. \frac{d^{j-1}}{d\tau^{j-1}} l_j(\tau) \right|_{\tau=b} \quad (\text{only valid for } j \geq 1)$$

When the number of components $m = 1$, the formulas for the approximate derivatives reduce to those of the classical ODE polynomials described in [14]. As is the case for classical ODE polynomials, the family of all additive ODE polynomials is large. Therefore, to simplify method construction, we focus on two special families that depend on significantly fewer free parameters.

4.1.1. Special families of partitioned ODE polynomials. We can generalize the Adams and BDF subfamilies (3.4) and (3.6) for the partitioned ODE polynomial (4.4). In particular, the following hold:

1. A *partitioned Adams ODE polynomial* has approximate derivatives

$$(4.5) \quad a_j(b) = \begin{cases} L_y(b), & j = 0, \\ \left. \frac{d^{j-1}}{d\tau^{j-1}} \sum_{k=1}^m L_f^{\{k\}}(\tau) \right|_{\tau=b}, & 1 \leq j \leq g, \end{cases}$$

where the following hold:

- $L_y(\tau) \approx y(t(\tau))$ is a Lagrange interpolating polynomial that interpolates only solution data in $D(r, s)$; specifically, for any number of distinct indices $\nu \in \{1, \dots, w\}$, $L_y(\tau_\nu) = y_\nu$.
- Each $L_f^{\{k\}}(\tau) \approx r f^{\{k\}}(t(\tau), y(t(\tau)))$ is a Lagrange interpolating polynomial of degree $g - 1$ that satisfies g data values pertaining to the k th derivative component; specifically, for g distinct indices $\nu \in \{1, \dots, w\}$, $L_f^{\{k\}}(\tau_\nu) = r f_\nu^{\{k\}}$.

Like the classical Adams ODE polynomial (4.6), all partitioned Adams ODE polynomials can be expressed in integral form as

$$(4.6) \quad p(\tau; b) = L_y(b) + \int_b^\tau \sum_{k=1}^m L_f^{\{k\}}(\xi) d\xi.$$

2. A *partitioned BDF ODE polynomial* has approximate derivatives that all satisfy

$$(4.7) \quad a_j(b) = \left. \frac{d^j}{d\tau^j} H_y(\tau) \right|_{\tau=b} \implies p(\tau; b) = H_y(\tau),$$

where $H_y(\tau) \approx y(t(\tau))$ is an interpolating polynomial of degree g that satisfies $g - 1$ solution values y_j , and whose derivative $H'_y(\tau)$ satisfies one total derivative F from Definition 4.2; specifically for g distinct indices $\nu \in \{1, \dots, w\}$, $H_y(\tau_\nu) = y_\nu$, and for one index $s \in \{1, \dots, w\}$, $H'_y(\tau_s) = F$.

If the number of components $m = 1$, then both formulas reduce to the classical Adams ODE polynomials (3.4) and the classical BDF ODE polynomials (3.6).

4.2. Additive polynomial block methods. Block methods [41] are multivalued integrators that advance a set (or “block”) of points at each timestep. If we take $s = 0$ in (3.8), then we obtain a PBM [14, 9, 11]. PBMs are simpler to derive compared to more general polynomial GLMs since all the ODE polynomials that determine the outputs are constructed using only input and output data. For this reason, PBMs were used to introduce both classical polynomial methods [14] and exponential polynomial methods [13]; in this paper we will use PBMs once more to introduce additive polynomial integrators.

An additive PBM depends on the parameters

$$\begin{aligned} q & \text{ number of inputs/outputs, } \{z_j\}_{j=1}^q \text{ nodes, } z_j \in \mathbb{C}, |z_j| \leq 1, \\ r & \text{ node radius, } r \geq 0, \quad \{b_j\}_{j=1}^q \text{ expansion points,} \\ \alpha & \text{ extrapolation factor} \end{aligned}$$

and can be written as

$$(4.8) \quad y_j^{[n+1]} = p_j(z_j + \alpha; b_j), \quad j = 1, \dots, q,$$

where each $p_j(\tau; b)$ is a partitioned ODE polynomial built from the partitioned ODE dataset

$$D(r, t_n) = \begin{cases} \text{inputs: } \left\{ \left(z_j, y_j^{[n]}, f_j^{\{1\}[n]}, \dots, f_j^{\{m\}[n]} \right) \right\}_{j=1}^q, \\ \text{outputs: } \left\{ \left(z_j + \alpha, y_j^{[n+1]}, f_j^{\{1\}[n+1]}, \dots, f_j^{\{m\}[n+1]} \right) \right\}_{j=1}^q. \end{cases}$$

Any additive PBM can be written in coefficient form as

$$(4.9) \quad \mathbf{y}^{[n+1]} = \mathbf{A}(\alpha) \mathbf{y}^{[n]} + r \sum_{k=1}^m \mathbf{B}^{\{k\}}(\alpha) \mathbf{f}^{\{k\}[n]} + r \sum_{k=1}^m \mathbf{C}^{\{k\}}(\alpha) \mathbf{f}^{\{k\}[n+1]},$$

where the matrices $\mathbf{A}(\alpha)$, $\mathbf{B}^{\{k\}}(\alpha)$, $\mathbf{C}^{\{k\}}(\alpha) \in \mathbb{R}^{q \times q}$, are of dimension $q \times q$, and the solution vectors and k th derivative component vectors are defined using $\mathbf{y}^{[n]} = [y_1^{[n]}, \dots, y_q^{[n]}]^T$ and $\mathbf{f}_j^{\{k\}[n]} = [f_1^{\{k\}[n]}, \dots, f_q^{\{k\}[n]}]^T$. Any additive polynomial PBM can be recast as an additive GLM whose parameters are parametrized in terms of the extrapolation parameter α . We show this in section SM4 for the doubly partitioned case when $m = 2$.

Last, additive PBMs are just one example of the more general class of additive polynomial GLMs, described in the following remark.

Remark 4.4 (additive polynomial GLMs). The formula for an additive polynomial method with s stages and q outputs is equivalent to (3.8) except that each $p_j(\tau; b)$ is now a partitioned ODE polynomial constructed from the partitioned ODE dataset that containing the method's input, output, and stage values. The family of additive polynomial GLMs is vast, and therefore we cannot provide a detailed exploration of such methods in this work. However, in section 5.2 we will see one example of an additive polynomial GLM that is constructed by composing multiple PBMs.

4.3. Linear stability. Linear stability analysis [30, Chapter IV, Section 2] characterizes the behavior of a time integrator applied to the Dahlquist test problem $y' = \lambda y$ and is essential for determining which types of problems cause instabilities. To study the linear stability properties of an additive integrator, we use the partitioned Dahlquist equation [6, 40, 31]

$$(4.10) \quad y' = \sum_{k=1}^m \lambda_k y, \quad y(0) = y_0,$$

where $\lambda_k y$ represents the k th component and $\lambda_k \in \mathbb{C}$. The nonlinear problem (4.1) reduces to the partitioned Dahlquist equation if all $f^{\{k\}}(t, y)$ are autonomous, diagonalizable linear operators that share the same eigenvectors. We note that this will rarely hold true in practice; nevertheless, this type of analysis is useful for studying and comparing the stability properties of additive methods.

When applied to (4.10), the additive polynomial integrator (4.9) reduces to the matrix iteration

$$(4.11) \quad \mathbf{y}^{[n+1]} = \mathbf{M}(z_1, z_2, \dots, z_m, \alpha) \mathbf{y}^{[n]},$$

where $\mathbf{y}_k^{[n]} = y_k^{[n]}$, $k = 1, \dots, q$, $\mathbf{M}: \mathbb{C}^m \times \mathbb{R}^+ \rightarrow \mathbb{C}^{q \times q}$, $z_j = h \lambda_j$, and h is the timestep. The stability region S is the subset of $\mathbb{C}^m \times \mathbb{R}^+$, where $\mathbf{M}(z_1, \dots, z_m, \alpha)$ is power bounded, so that

$$(4.12) \quad S = \left\{ \mathbf{z} \in \mathbb{C}^m, \alpha \in \mathbb{R}^+ \mid \sup_{n \in \mathbb{N}} \|\mathbf{M}(\mathbf{z}_1, \dots, \mathbf{z}_m, \alpha)^n\| < \infty \right\}.$$

Therefore, the method is stable if the eigenvalues of $\mathbf{M}(z_1, \dots, z_m, \alpha)$ lie inside the closed unit disk, and any eigenvalues of magnitude one are nondefective. In section 5 we will study the stability of additive polynomial integrators with $m = 2$ and define several two-dimensional subsets of S that are useful for visualizing the stability region.

5. Constructing FIMEX polynomial integrators based on Radau IIA. The additive polynomial framework enables the construction of additive PBMs for the general partitioned equation (4.1). However, in this introductory work we will focus on constructing FIMEX methods for the simpler doubly partitioned equation (2.1), where $f^{\{1\}}$ is treated implicitly and $f^{\{2\}}$ is treated explicitly. In particular, we will

introduce a family of arbitrary-order FIMEX integrators that are based on the fully implicit Radau IIA methods described in section 3.1.

To introduce a FIMEX method with an implicit part that is equivalent to Radau IIA, we start by reusing the nodes (3.11), where we assume that $q \geq 2$. As in section 3.1, we construct a PBM whose outputs are all computed using a single Adams ODE polynomial $p(\tau; b)$ such that

$$(5.1) \quad y_j^{[n+1]} = p(z_j + \alpha; b_j), \quad j = 1, \dots, q.$$

Since we are considering an additive integrator for two component systems, the Adams polynomial is now doubly partitioned. To obtain an IMEX integrator, we must treat the first component implicitly and the second component explicitly, that is,

$$(5.2) \quad p(\tau; b) = L_y(b) + \underbrace{\int_b^\tau L_f^{\{1\}}(\xi) d\xi}_{\text{implicit approx.}} + \underbrace{\int_b^\tau L_f^{\{2\}}(\xi) d\xi}_{\text{explicit approx.}}.$$

If we want an implicit component that is equivalent to Radau IIA, then we must select $L_y(\tau)$, $L_f^{\{1\}}(\tau)$, b , and α identically to what was done in section 3.1, namely, the following hold:

- $L_y(\tau)$ is the polynomial of degree $q - 1$ that interpolates all q inputs, such that $L_y(z_j) = y_j^{[n]}$, $j = 1, \dots, q$.
- $L_f^{\{1\}}(\tau)$ is the polynomial of degree $q - 2$ that interpolates the last $q - 1$ output component derivatives $f_j^{\{1\}[n+1]}$ such that $L_f^{\{1\}}(z_j + \alpha) = r f_j^{\{1\}[n+1]}$, $j = 2, \dots, q$.
- The extrapolation factor is $\alpha = 2$ and the lower integration bound is $b = 1$.

To form the explicit approximation, we now make use of the input derivative components $f_j^{\{2\}[n]}$. We propose two different strategies for selecting $L_f^{\{2\}}(\tau)$:

1. Use only the derivative components at the input Radau nodes; let $L_f^{\{2\}}(\tau)$ be a polynomial of order $q - 2$ that satisfies

$$(5.3) \quad L_f^{\{2\}}(z_j) = r f_j^{\{2\}[n]}, \quad j = 2, \dots, q.$$

2. Use all input derivative components; let $L_f^{\{2\}}(\tau)$ be a polynomial of order $q - 1$ that satisfies

$$(5.4) \quad L_f^{\{2\}}(z_j) = r f_j^{\{2\}[n]}, \quad j = 1, \dots, q.$$

Both choices for $L_f^{\{2\}}(\tau)$ lead to the multivalued FIMEX method

$$(5.5) \quad y_j^{[n+1]} = y_j^{[n]} + \int_1^{z_j+2} L_f^{\{1\}}(\xi) + L_f^{\{2\}}(\xi) d\xi, \quad j = 1, \dots, q,$$

where the implicit component is equivalent to Radau IIA; notice that if $f^{\{2\}}(t, y) = 0$, then $L_f^{\{2\}}(\tau) = 0$ and (5.5) is equivalent to (3.12). The order-of-accuracy of (5.5) is equal to the minimum order-of-accuracy of the implicit and explicit component. Because the implicit component is equivalent to Radau IIA with $q - 1$ stages, its order-of-accuracy is $2q - 3$. Depending on whether we choose the polynomial (5.3) or (5.4) the explicit component, respectively, has an order of $q - 1$ or q . Therefore, for all $q > 2$, the explicit component determines the overall accuracy and the higher-order

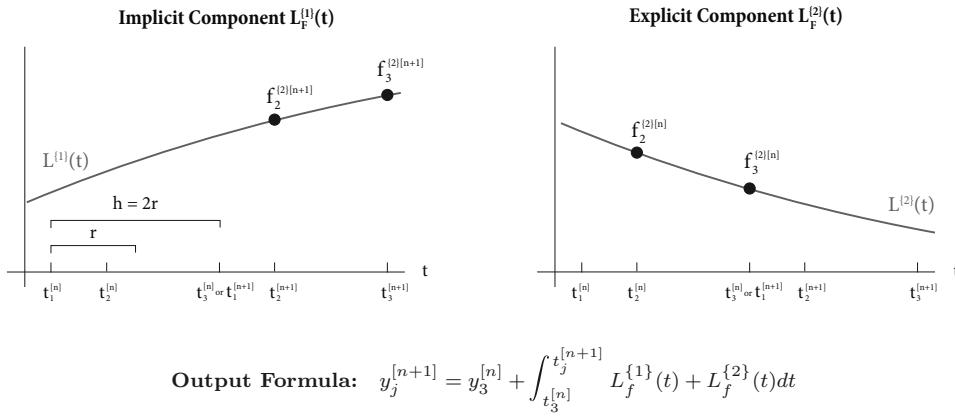


FIG. 3. An illustration depicting the polynomials in global time t that determine the outputs of the polynomial FIMEX-Radau method (5.5) with $q=3$ and nodes $\{z_j\} = \{-1, -1/3, 1\}$. The nodes $t_j^{[n]}$ are the nodes z_j in global coordinates at the n th timestep such that $t_j^{[n]} = t_n + rz_j$.

polynomial is desirable. However, in section 5.3 we will see that the higher-order polynomial also leads to inferior stability properties.

From here on, we will refer to the FIMEX method with (5.3) as FIMEX-Radau and the FIMEX method with (5.4) as FIMEX-Radau*. In Figure 3 we present an illustration of the polynomials $L_f^{(1)}(\tau)$ and $L_f^{(2)}(\tau)$ for a FIMEX-Radau method with $q=3$, and below we also show the coefficients for the method:

$$(5.6) \quad \begin{aligned} y_1^{[n+1]} &= y_3^{[n]}, \\ y_2^{[n+1]} &= y_3^{[n]} + \frac{5}{6}rf_2^{(1)[n+1]} - \frac{1}{6}rf_3^{(1)[n+1]} - \frac{1}{6}rf_2^{(2)[n]} + \frac{5}{6}rf_3^{(2)[n]}, \\ y_3^{[n+1]} &= y_3^{[n]} + \frac{3}{2}rf_2^{(1)[n+1]} + \frac{1}{2}rf_3^{(1)[n+1]} - \frac{3}{2}rf_2^{(2)[n]} + \frac{7}{2}rf_3^{(2)[n]}. \end{aligned}$$

Notice that if $f^{(2)}(t, y) = 0$, then (5.6) is equivalent to (3.13).

Any FIMEX-Radau or FIMEX-Radau* method with q nodes can be written in the coefficient form

$$(5.7) \quad \mathbf{y}^{[n+1]} = \mathbf{A}\mathbf{y}^{[n]} + r\mathbf{B}^{(1)}\mathbf{f}^{(1)[n+1]} + r\mathbf{B}^{(2)}\mathbf{f}^{(2)[n]},$$

where $\mathbf{A}, \mathbf{B}^{(1)}, \mathbf{B}^{(2)} \in \mathbb{R}^{q \times q}$. In supplementary materials section SM1.1 we provide the coefficients for $q=2, 3, 4$, along with a MATLAB script for initializing coefficients for larger q . When $q=2$, the FIMEX-Radau method is equivalent to the IMEX-Euler method (2.3), while FIMEX-Radau* treats the implicit component with backward Euler and the explicit component with second-order Adams–Bashforth. For $q > 2$, the nonlinear implicit equations that arise in both FIMEX-Radau methods (5.7) are analogous to those that arise in standard Radau IIA integration, with a modified right-hand side derived from the explicit component (i.e., linear and nonlinear solvers developed for fully implicit RK [18, 38, 26, 32, 39, 45, 44] naturally apply to FIMEX-Radau).

Note that the method construction described here also applies to different node sets. For example, if we had selected Legendre nodes, then the implicit component would be equivalent to that of the fully implicit Gauss methods. However, Radau nodes have the advantage that they lead to an L-stable implicit component. Last,

this method construction shares many similarities with the exponential PBM methods based on Legendre nodes from [13]. Loosely speaking, we have traded the exponential for a fully implicit approximation and switched from Legendre nodes to Radau nodes.

5.1. A FIMEX-Radau iterator for obtaining initial conditions. PBMs with $q > 1$, including (5.5), require multiple inputs at the first timestep. Though a one-step method can be used to compute these solution values, it may not always be possible to match the order of a starting method with that of the PBM. In [13] we proposed to use an iterator (a polynomial method with $\alpha = 0$ described in section 3.2) to compute the initial conditions of exponential PBMs. The iterator used a discrete exponential Picard iteration to improve the accuracy of an approximate solution. We can reuse the same idea in the context of additive integrators. In fact, it is always possible to construct a polynomial iterator that improves the accuracy of solution values relative to one or more highly accurate inputs [9, Chap. 6.2].

We can obtain a fully implicit iterator by replacing $\alpha = 2$ with $\alpha = 0$ in (5.5) so that the upper integration bound is just z_j . However, this iterator steps backward in time, which introduces additional complications. By modifying the method construction so that $b_j = -1$, we can avoid integrating backward in time. This leads to the iterator

$$(5.8) \quad y_j^{[n+1]} = y_1^{[n]} + \int_{-1}^{z_j} L_f^{\{1\}}(s) + L_f^{\{2\}}(s) ds, \quad j = 1 \dots, q,$$

which uses the Radau IIA coefficients for both the implicit and explicit parts. For example, when $q = 3$, we obtain the method

$$(5.9) \quad \begin{aligned} y_1^{[n+1]} &= y_1^{[n]}, \\ y_2^{[n+1]} &= y_1^{[n]} + \frac{5}{6}rf_2^{\{1\}}[n+1] - \frac{1}{6}rf_3^{\{1\}}[n+1] + \frac{5}{6}rf_2^{\{2\}}[n] - \frac{1}{6}rf_3^{\{2\}}[n], \\ y_3^{[n+1]} &= y_1^{[n]} + \frac{3}{2}rf_2^{\{1\}}[n+1] + \frac{1}{2}rf_3^{\{1\}}[n+1] + \frac{3}{2}rf_2^{\{2\}}[n] + \frac{1}{2}rf_3^{\{2\}}[n]. \end{aligned}$$

Note that both the implicit and explicit coefficients in the iterator (5.9) are identical to the implicit coefficients in the propagator (5.6). If $y_1^{[n]}$ is of order $2q - 3$ or higher, then every application of the iterator (5.8) will improve the order-of-accuracy of all other solution values by one, up to a maximum order of $2q - 3$. Moreover, when a repeatedly applied iteration converges, then the PBM outputs are equivalent to the output and stage values of a fully implicit Radau IIA method with $q - 1$ stages.

We can therefore use (5.8) to compute initial conditions in the following way. We first consider an iterator with quadrature points on the interval $[0, 2]$, instead of $[-1, 1]$. The method has identical coefficients, but the first input node at $z_1 = 0$ is now located at $t = t_0$, where the initial condition for (1.1) is provided (i.e., $y_1^{[0]} = y(rz_1 + t_0) = y_0$). Next we obtain a zeroth-order estimate for the solution at all the input nodes by temporarily assuming a constant solution such that $y_j^{[n]} = y_0$. Last, we repeatedly apply the iterator to improve the accuracy of the zeroth-order estimate. This procedure provides an accurate initial solution at the times $t = t_0 + rz_j$ for $j = 1, \dots, q$ and can be written abstractly as

$$(5.10) \quad \mathbf{y}^{[0]} = M^\kappa(\mathbf{c}), \quad \mathbf{c}_j = y_0, \quad j = 1, \dots, q,$$

where \mathbf{c} is the initial zeroth-order approximation, and $M^\kappa(\mathbf{c})$ is the iterator method applied the κ times to initial condition \mathbf{c} such that $M^\kappa(\mathbf{c}) = M(M(\dots(M(\mathbf{c}))))$. For a FIMEX-Radau or FIMEX-Radau* method, the iterator should be, respectively, applied a minimum of $\kappa = q - 1$ and $\kappa = q$ times at the first step to match the order of the explicit component.

5.2. Composite FIMEX-Radau methods. The method (5.5), with $L_f^{\{2\}}(\tau)$ selected using (5.3), pairs a fully implicit integrator of order $2q - 3$ with an explicit integrator of order $q - 1$. Therefore, the combined order is limited to $q - 1$. Using (5.4) improves the order by one, but the imbalance in accuracy between the implicit and explicit components remains. One strategy for increasing accuracy further is to use a composite method that first advances the timestep with the propagator (5.6) and then corrects the output of the propagator κ times using the iterator (5.9). This idea of composite PBMs was introduced in [13] for exponential integrators and shares many similarities with spectral deferred correction methods [24, 37, 10]. In the following sections we show how the same idea leads to composite additive integrators with improved stability and accuracy properties.

The composite method can be written abstractly as

$$(5.11) \quad \mathbf{y}^{[n+1]} = M^\kappa(P(\mathbf{y}^{[n]})),$$

when $P(\cdot)$ denotes the propagator (5.5), where $L_f^{\{2\}}$ is selected using (5.3) or (5.4), and $M^\kappa(\cdot)$ denotes κ applications of the iterator (5.8). We will refer to this composite method as FIMEX-Radau(q, κ) if P is (5.3) and FIMEX-Radau* (q, κ) if P is (5.4); q is the number of nodes and κ is the number of iterator applications. Since each application of the iterator improves the accuracy order by one, the associated order-of-accuracy for these methods is

$$(5.12) \quad \text{FIMEX-Radau}(q, \kappa) : \min(2q - 3, q - 1 + \kappa),$$

$$(5.13) \quad \text{FIMEX-Radau}^*(q, \kappa) : \min(2q - 3, q + \kappa).$$

Last, it is important to note that when using iterative methods to solve the implicit equations (as in numerical PDEs), the implicit solve for the iterator will typically be significantly faster than the propagator, due to a very good initial guess. That is, rather than advance a set of solutions forward in time by h like a propagator, an iterator simply increases the accuracy of the *current* solutions by one.

5.3. Linear stability. We now discuss the linear stability properties of FIMEX-Radau* and FIMEX-Radau. If we select the splitting (2.4), then the stability region of the additive integrator is equivalent to that of the Radau IIA method used for the implicit component. Because all Radau IIA methods are L-stable [30, Chapter IV, Section 5], both FIMEX methods will have excellent stability. For the remaining splittings (2.4)–(2.6) we must instead consider the four dimensional stability region

$$(5.14) \quad S = \{z_1, z_2 \in \mathbb{C} \mid \rho(\mathbf{M}(z_1, z_2, \alpha = 2)) \leq 1\},$$

where $\rho(\cdot)$ denotes the spectral radius. Due to the high-dimensionality of S , it is convenient to consider two-dimensional slices that are formed by fixing z_1 :

$$(5.15) \quad S(z_1) = \{z_2 \in \mathbb{C} \mid \rho(\mathbf{M}(z_1, z_2, \alpha = 2)) \leq 1\}.$$

The magnitude and argument of the complex number z_1 , respectively, determine the stiffness of the implicit component and the degree of diffusion.

In the PDE setting, $\arg(z_1) = \pi/2$ approximately represents a skew-symmetric advection discretization, while $\arg(z_1) = \pi$ approximately represents a symmetric positive-definite diffusion discretization. An intermediate value, $\arg(z_1) \in (\pi/2, \pi)$, approximately represents a mix of advection and diffusion. However, the stability region $S(z_1)$ will not be symmetric along the real z_1 axis (i.e., $S(z_1) \neq S(z_1^*)$). In

contrast all real-valued discretizations of PDEs have spectrums that are symmetric about the imaginary axis. We therefore consider the stricter stability region

$$(5.16) \quad \hat{S}(z_1) = \{z_2 \in \mathbb{C} \mid \max [\rho(\mathbf{M}(z_1, z_2, \alpha = 2)), \rho(\mathbf{M}(z_1^*, z_2, \alpha = 2))] < 1\}.$$

Last, it is also useful to consider the region

$$(5.17) \quad \tilde{S}(\theta) = \{z_2 \in \mathbb{C}, \theta \in \mathbb{R}^+ \mid \max [\rho(\mathbf{M}(\gamma e^{i\theta}, z_2, \alpha = 2)), \rho(\mathbf{M}(\gamma e^{i\theta}, z_2, \alpha = 2))] \},$$

which contains all the z_2 values that ensure stability for any $z_1 = \gamma e^{i\omega}$ in the wedge $\gamma \geq 0$ and $\theta \leq \omega \leq 2\pi - \theta$.

To present important subsets of the full stability region S we overlay multiple contours of the $\hat{S}(z_1)$, where we take $\arg(z_1) \in \{\pi, \frac{3\pi}{4}, \frac{\pi}{2}\}$ and $|z_1| \in \{0, 3, 6\}$. These choices of $\arg(z_1)$, respectively, approximate an implicit linear component with diffusion, a mix of diffusion and oscillation, and pure oscillation. In Figure 4 we show these stability regions for the composite method FIMEX-Radau*(4, κ) from (5.11) with $\kappa = 0, 1, 2$. When $\kappa = 0$ the FIMEX-Radau*(4, 0) method is equivalent to the FIMEX-Radau* method (5.5) with $q = 4$.

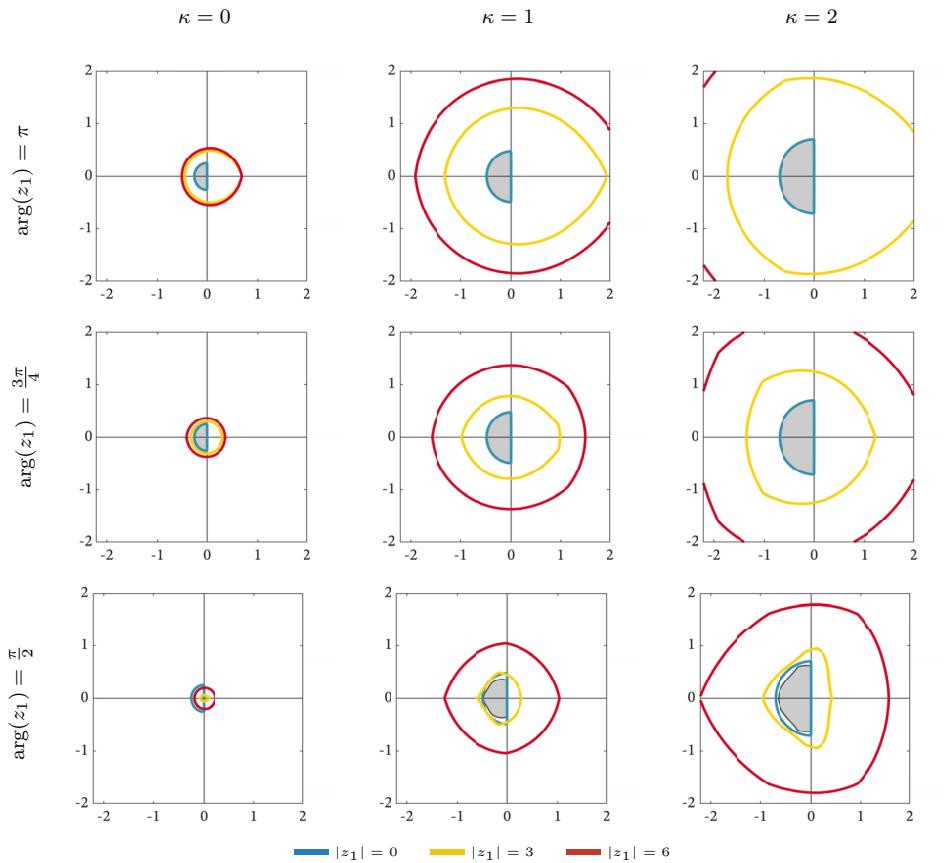


FIG. 4. Stability regions for composite FIMEX-Radau*($q = 4, \kappa$). Each colored contour represents a different stability region $\hat{S}(z_1)$ defined in (5.16). The gray region is the stability region $\tilde{S}(\arg(z_1))$ described in (5.17). (Color images are available online.)

The stability regions for FIMEX-Radau* are largest for the diffusive case with $\arg(z_1) = \pi$ and smallest for the oscillatory case with $\arg(z_1) = \pi/2$. For each of the three values of $\arg(z_1)$, the FIMEX-Radau*(4, κ) stability regions always enclose the imaginary axis. Despite the fact that FIMEX-Radau* is equipped with a fully implicit propagator, the stability regions grow slowly in $|z_1|$ and are very small when $\arg(z_1) = \frac{\pi}{2}$. However, even a single application of the iterator leads to significantly larger linear stability regions for all three values of $\arg(z_1)$. Last, in Figure SM1 of the supplemental materials we also show the equivalent stability regions for the composite FIMEX-Radau method; overall we see that using a lower-order polynomial approximation for the explicit component leads to improved stability.

6. Numerical experiments. We investigate the efficiency of FIMEX-Radau and FIMEX-Radau* methods by conducting three types of numerical experiments. First, we study the performance of the integrators on PDEs with periodic domains where inverting the implicit component is trivial. Then, we compare both method families on a finite-element discretization of a nonperiodic problem, where solving the fully implicit system requires special care [26, 32, 39, 45, 44]. Last, we numerically investigate order-reduction on the singularly perturbed Van der Pol equation.

For certain spatial discretizations, such as spectral methods, and/or solvers that are difficult to parallelize, the FIMEX-Radau methods allow for certain types of parallelization that are not feasible using existing IMEX-RK methods. For example, when considering periodic domains with Fourier discretizations, the implicit component is trivial to invert, and the majority of the cost is due to the nonlinear function evaluations. In subsection 6.1 we demonstrate how parallelization of nonlinear function evaluations can be applied to obtain highly accurate solutions in significantly less time. More generally, there are situations such as when solver performance degrades in parallel or when finite element construction and nonlinear function evaluations can be very expensive, where the additional parallelization provided by these FIMEX-Radau can be exploited.

Conversely, with finite element discretizations and scalable solvers, spatial parallelism is often very effective. In subsection 6.3 we investigate the efficiency of FIMEX-Radau methods with traditional spatial parallelism on a problem where function evaluations are relatively cheap and time parallelism is not needed.

6.1. PDEs with periodic boundary conditions. In this experiment we evaluate the accuracy and efficiency of composite FIMEX-Radau* methods (5.11) on the dispersive Korteweg–De Vries (KDV) equation with periodic boundary conditions. For comparison we also include results for IMEX-RK integrators of orders one to four; specifically the (1,1,1) and (2,3,2) methods from [5, sects. 2.1 and 2.5] and the ARK3(2)4L[2]SA and ARK4(3)6L[2]SA from [33]. The numerical experiment is identical to one found in [13, 10] and is only briefly described below:

The KDV equation from [10, 13]

$$\frac{\partial u}{\partial t} = - \left[\delta \frac{\partial^3 u}{\partial x^3} + \frac{1}{2} \frac{\partial}{\partial x} (u^2) \right], \quad u(x, t=0) = \cos(\pi x), \quad x \in [0, 2],$$

where $\delta = 0.022$. This equation is integrated to time $t = 3.6/\pi$ using a 512 point Fourier spectral discretization.

6.1.1. Implementation details. We solve the KDV equation in Fourier space where the initial value problem has the form $\mathbf{y}' = \mathbf{Ly} + \mathbf{N}(t, \mathbf{y})$, where \mathbf{L} is an $N \times N$

diagonal matrix that includes the discretized linear differential operator. For dealiasing we apply the standard 3/2 rule that zeros out the top third of the spectrum. The reference solution \mathbf{y}_{ref} is computed using a 32nd-order exponential spectral deferred correction method [10], and the relative error in our convergence and efficiency studies is defined as $\|\mathbf{y} - \mathbf{y}_{\text{ref}}\|_{\infty}/\|\mathbf{y}_{\text{ref}}\|$.

Since the matrix \mathbf{L} is diagonal, the implicit solve for a fully implicit propagator or iterator (5.7) amounts to inverting N decoupled $q \times q$ systems of the form $(\mathbf{I} - r\mathbf{L}_{kk}\mathbf{B}^{\{2\}})\mathbf{x} = \mathbf{b}$ for $k = 1, \dots, N$. To save computational time the matrix inverses for propagators and iterators are precomputed and stored at the first timestep so that an implicit solve can be done using $N, q \times q$ matrix multiplications. Since this operation is cheap, the dominant computational cost is due to the nonlinear function evaluations.

FIMEX-Radau integrators benefit from time parallelism since the nonlinear function evaluations and output computations can be computed simultaneously. To quantify the benefits of parallelism, we created both serial Fortran implementations of the integrators and parallel Fortran implementations using OpenMP. A FIMEX-Radau method with q nodes requires q independent function evaluations that can be computed simultaneously using q shared or distributed memory processes/threads. In contrast an IMEX-RK method requires sequential evaluation of the nonlinear term at each stage, and therefore this type of parallelization is not possible. However, all methods can benefit from spatial parallelization where the FFT evaluations are computed using multiple threads. For the purposes of this experiment we will only investigate time parallelism and compute all FFT evaluations using one thread.

Our Fortran code can be found in [12], and all the timing results presented in this subsection were produced on a 14 core, 2.0 Ghz Intel Xeon E5-2683 v3 with hyper-threading enabled.

6.1.2. Results and discussion. In Figure 5 we present results for the KDV equation solved using composite FIMEX-Radau $^*(q, 2)$ methods with $q = 2, 3, 4, 5$.

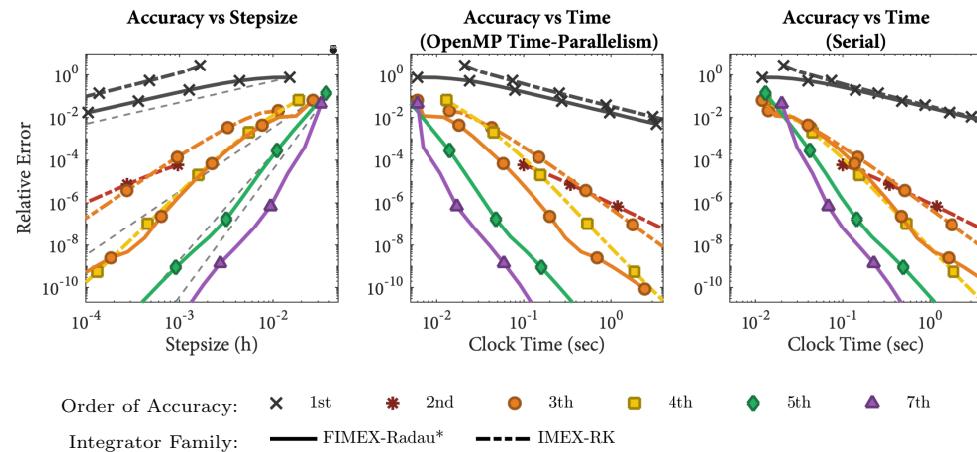


FIG. 5. Accuracy and precision diagrams comparing composite FIMEX-Radau $^*(q, 2)$ with $q = 2, 3, 4, 5$ to IMEX-RK on the KDV equation. All curves start at the first stable timestep. The parallel and serial time plots, respectively, show run times for composite FIMEX-Radau $^*(q, 2)$ methods with and without OpenMP parallelization. The dashed lines of increasing slope in the accuracy plot, respectively, correspond to first-, third-, fifth-, and seventh-order convergence, which are the expected orders of convergence (5.13) for FIMEX-Radau * .

We select $\kappa = 2$ since the linear stability regions from Figure 4 revealed that FIMEX-Radau*($q, 0$) have comparatively poor stability for nondiffusive equations. Overall the FIMEX-Radau*($q, 2$) integrators performed excellently, with the serial, seventh-order FIMEX-Radau*($5, 2$) outperforming all the other methods for any accuracy below 10^{-2} . When using OpenMP parallelism with five threads, the FIMEX-Radau*($5, 2$) is the most efficient method across all accuracies and is capable of obtaining the solution with a relative error of 10^{-2} approximately two times faster than any RK method. For more accurate solutions with errors below 10^{-4} , the difference was even more significant and FIMEX-Radau*($5, 2$) was able to obtain the solution approximately thirty times faster than IMEX-RK4. For lower-order methods, the third-order FIMEX-Radau*($3, 2$) method exhibited increased fourth-order convergence throughout most of the stepsize range. This enabled the method to be more efficient than the IMEX-RK4 method, despite only requiring the storage of three solution vectors instead of the five needed by the RK method. Last, if very inaccurate solutions are sufficient, then the FIMEX-Radau*($2, 1$) method was both more efficient and more stable than the IMEX-RK1 (or equivalently FIMEX-Radau($2, 0$)) method.

6.2. Numerically investigating order-reduction. In this experiment we numerically investigate order-reduction for the composite FIMEX-Radau and FIMEX-Radau* method (5.11) by solving the Van der Pol equation [30, p. 403]

$$(6.1) \quad \begin{aligned} y'_1 &= y_2, & y_1(0) &= 2, \\ y'_2 &= \frac{(1 - y_1^2)y_2 - y_1}{\epsilon}, & y_2(0) &= -\frac{2}{3} + \frac{10}{81}\epsilon - \frac{292}{2187}\epsilon^2 - \frac{1814}{19683}\epsilon^3, \end{aligned}$$

integrated to time $t = 0.5$. We consider ϵ values ranging from $\epsilon = 1$ to $\epsilon = 10^{-8}$. When the stiffness parameter ϵ is small, this equation is known to cause order-reduction for IMEX-RK methods [5, 7, 33, 35, 31, 34].

We integrate the Van der Pol equation using composite methods (5.11), where the propagator uses (5.3) or (5.4), $q \in \{3, 4, 5\}$, and $\kappa \in \{0, 1, 2\}$. The equation's right-hand side is split in two different ways. A *semi-implicit* splitting treats the first component explicitly and the second component implicitly; this same splitting was used in [31, 5, 7, 35, 33, 34]. We also consider the semilinear splitting (2.8), where J_n is the exact Jacobian of the right-hand side at the n th timestep.

To verify FIMEX-Radau(q, κ) and FIMEX-Radau*(q, κ) methods, we estimate their convergence rates as a function of ϵ using the approach proposed in [34]. Specifically, we solve (6.1) using 30 logarithmically spaced stepsizes ranging from $h = 0.25$ to $h = 10^{-4}$ and then compute a linear least-squares fit of $\log(\text{error})$ versus $\log(\text{stepsize})$. Error for an approximate solution \mathbf{y} is measured using $\|\mathbf{y} - \mathbf{y}_{\text{ref}}\|_\infty$, where \mathbf{y}_{ref} was computed using the MATLAB `ode15s` integrator. For the semi-implicit splitting we set the tolerance of our Newton iteration to 10^{-12} and use the MATLAB backslash function to solve the associated linear systems. For the linearly implicit system we also use the MATLAB backslash function to solve the linear systems at each timestep.

In Figure 6 we show convergence rate plots and convergence diagrams for the FIMEX-Radau*(q, κ) methods with each splitting. Convergence diagrams for different ϵ values are shown in Figure SM2 of the supplemental materials. We also show a convergence rate plot and convergence diagrams for the FIMEX-Radau(q, κ) in Figure SM3. There are several important points regarding the results:

- The FIMEX-Radau* and FIMEX-Radau methods are stable across the full range of stepsizes for both splittings. The linearly implicit splitting does not require a nonlinear solve at each step, leading to improved computational

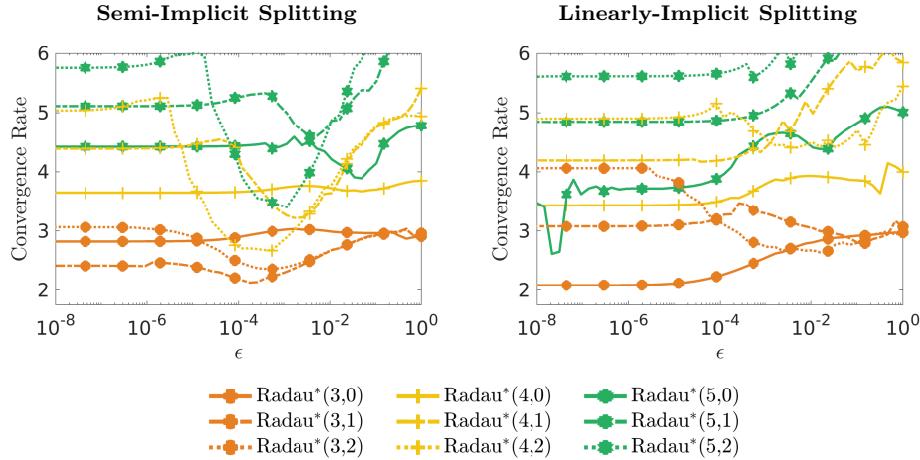


FIG. 6. Approximate overall convergence rates for the Van der Pol equation (6.1) as a function of ϵ for the FIMEX-Radau* method with a semi-implicit splitting or a linearly implicit splitting.

efficiency. In contrast, all the IMEX-RK methods we tested from [34, 33] required substantially smaller timesteps to remain stable with the linearly implicit splitting (see the supplemental Figure SM4).

- For small ϵ , each application of the iterator raises the order of convergence by approximately one. When using a semi-implicit splitting, we see a significantly decreased convergence rate for Radau($q,2$) and Radau*($q,2$) methods when ϵ is between 10^{-2} and 10^{-4} . However the extra iteration improves accuracy significantly. For example, from the convergence diagrams for $\epsilon = 10^{-3}$ in Figures SM2 and SM3, we see that methods with $\kappa = 2$ are always more accurate than those with $\kappa = 1$ or $\kappa = 0$, despite their reduced convergence rates.
- The estimated convergence rates are consistently smaller than the expected orders-of-accuracy (5.12) and (5.13). This is due to minor order-reduction at stepsizes $h \geq 0.1$ (less than five total timesteps). By including this data in our least-squares fits we see lower overall convergence rates. If we only consider $h < 0.1$, then we would observe expected convergence (for example, we can see that the convergence curves for $h < 0.1$ match the dashed convergence-order lines in the diagrams in Figures SM2 and SM3).
- The convergence rates for FIMEX-Radau*(3,2) and FIMEX-Radau(3,2) with a linearly implicit splitting are unusually high. This is due to rapid convergence at coarse timesteps that leads to higher than expected overall convergence. Nevertheless, for sufficiently small stepsizes, the convergence limits to third-order (e.g., see the convergence diagram for $\epsilon = 10^{-3}$ in Figures SM2 and SM3).
- The FIMEX-Radau*(5,0) method with a linearly implicit splitting did not converge at large timesteps (though it was stable). Increasing κ to one resolves the issue.

6.3. DG advection-diffusion-reaction. Here we consider the time-dependent advection-diffusion-reaction equation

$$(6.2) \quad u_t + \nabla \cdot (\beta u - \epsilon \nabla u) + \gamma u^2 = f,$$

where $\beta(x, y) := (1, 1)^T$ is the velocity field and ϵ and γ constant diffusion and reaction coefficients.

We discretize (6.2) in space using discontinuous Galerkin finite elements over the spatial domain $\Omega = [0, 1] \times [0, 1]$. Dirichlet boundary conditions are weakly enforced on $\partial\Omega$, advection terms are upwinded [20], and diffusion terms are treated with the symmetric interior penalty method [3, 4]. Let V_h be the DG finite element space consisting of piecewise polynomials of degree p defined locally on elements of the spatial mesh \mathcal{T} . The resulting finite element problem is to find $u_h \in V_h$ such that, for all $v_h \in V_h$,

$$\begin{aligned} & \int_{\Omega} \partial_t(u_h)v_h \, dx - \int_{\Omega} u_h \beta \cdot \nabla_h v_h \, dx + \int_{\Gamma} \widehat{u_h} \beta \cdot [\![v_h]\!] \, ds + \int_{\Omega} \nabla_h u_h \cdot \nabla_h v_h \, dx + \int_{\Omega} \gamma u_h v_h^2 \, dx \\ & - \int_{\Gamma} \{\nabla_h u_h\} \cdot [\![v_h]\!] \, ds - \int_{\Gamma} \{\nabla_h v_h\} \cdot [\![u_h]\!] \, ds + \int_{\Gamma} \sigma [\![u_h]\!] \cdot [\![v_h]\!] \, ds = \int_{\Omega} f v_h \, dx, \end{aligned}$$

where ∇_h is the broken gradient, Γ denotes the skeleton of the mesh, $\{\cdot\}$ and $[\![\cdot]\!]$ denote the average and jump operators, respectively, and $\widehat{u_h}$ is used to denote the upwind numerical flux. The parameter σ is the *interior penalty parameter*, important for a stable discretization [4], which we set to $\sigma = (p+1)^2/h$.

The discretization is implemented in the MFEM finite element library [2]. Classical algebraic multigrid (AMG) in the hypre library [25] (BoomerAMG) is used to solve the implicit equations, with Falgout coarsening, classical interpolation, a strength tolerance of 0.2, and hybrid parallel Gauss–Seidel relaxation. For the Radau schemes, we must solve an implicit system analogous to that which arises in fully implicit RK methods. We wrap the AMG solver with the block preconditioning for fully implicit RK methods developed in [45, 44] to solve the Radau stage equations. We use fourth-order elements in space, with mesh spacing $h_x \approx 0.0078$ leading to expected spatial accuracy $\sim \mathcal{O}(10^{-9})$. We consider two problems and splittings:

1. No reaction: treat advection explicitly and diffusion and source term implicitly (subsection 6.3.1).
2. $\gamma = 10$: treat nonlinear reaction explicitly and advection-diffusion implicitly (subsection 6.3.2).

All simulations are run on 4 dual socket Intel Xeon Gold 6152 22-core processors (i.e., 44 cores/node).

6.3.1. Advection-diffusion. This section considers no reaction ($\gamma = 0$) and an additive splitting of (6.2) treating the advection explicitly and diffusion implicitly. We choose the forcing function $f(x, y, t)$ such that the analytical solution is given by $u_*(x, y, t) = \sin(2\pi x(1-y)(1+2t)) \sin(2\pi y(1-x)(1+2t))$. Here we solve all linear systems to a relative residual tolerance of 10^{-12} . All simulations are run to final (simulation) time $t_f = 2$ with 120 total MPI processes.

Results in this section indicate the superior stability and accuracy of FIMEX-Radau* methods over traditional IMEX-RK. However, results are somewhat artificial in this setting in that (i) there exist fast AMG solvers for advection-diffusion, even in the advection-dominated regime [36, 42], and (ii) due to stability issues with the additive splitting of this discretization shown below, even classical AMG in hypre [25] works well as-is applied to the advection-diffusion ratios considered here. That is, although FIMEX-Radau* methods show advantages over classical IMEX-RK, in practice one would probably treat this particular problem and discretization implicitly. Other related advection-diffusion equations, however, such as a DG-BDM discretization of incompressible Navier–Stokes, can be very difficult to solve implicitly

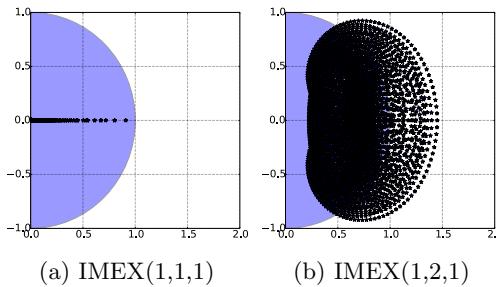


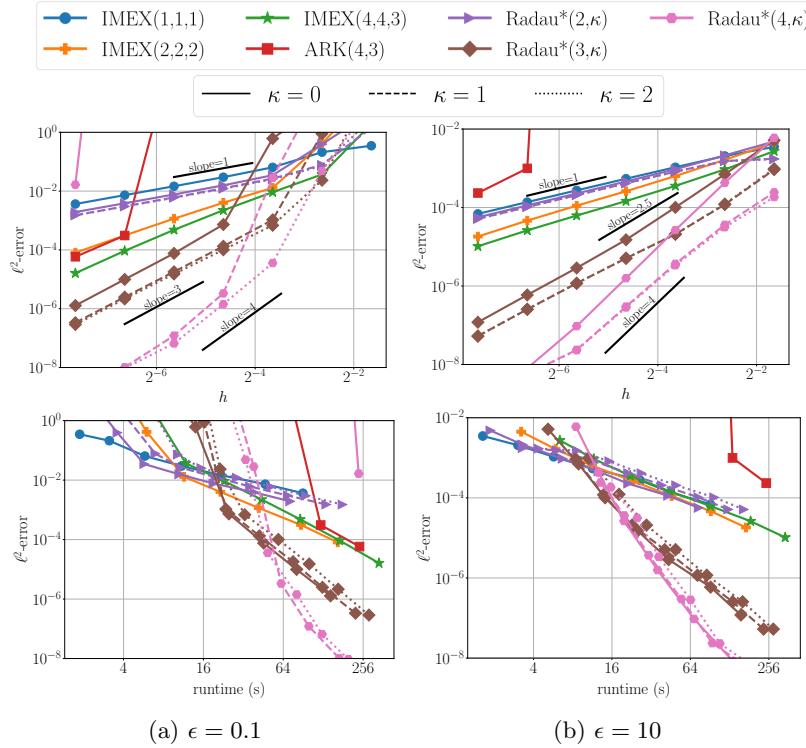
FIG. 7. Eigenvalues of propagation operator for two first-order IMEX-RK schemes from [5], with timestep $h = 0.004876$ given by the advective stability limit. The unit circle is shaded; to ensure stability, eigenvalues must be inside of the shaded region.

for nontrivial advection, so it is possible FIMEX-Radau* methods have a use in such advection-diffusion problems, where implicit treatment is difficult or not viable (note that incompressible Navier–Stokes is an index-2 DAE, which introduces other complications in terms of accuracy; analyzing FIMEX-Radau* methods in the DAE setting is ongoing work).

First note that the stability of IMEX integration for this discretization is challenging. Consider two variations in first-order IMEX-RK presented in [5], the IMEX(1,1,1) and IMEX(1,2,1) schemes. Note that a necessary condition for stability is having the eigenvalues of the propagation operator bounded < 1 in magnitude. Thus, we consider mesh spacing $h_x \approx 0.03$ and directly construct the propagation operators for the IMEX(1,1,1)- and IMEX(1,2,1)-schemes. Letting $\epsilon = 1$ and $h = 0.004876$ be the (numerically determined) forward Euler advective stability limit, eigenvalues for the two propagation operators are shown in Figure 7. Note that for this moderately small timestep, RK(1,2,1) is very unstable; numerical tests confirm this in practice, with the solution rapidly growing by orders of magnitude. In fact, many of the IMEX-RK schemes proposed in [5] suffer from similar instabilities with this problem and discretization. The *only* schemes from [5] that have demonstrated reasonable stability here are schemes in which *both* the implicit and explicit method are stiffly accurate. Moving forward, we only consider these schemes, namely, IMEX(1,1,1), RK(2,2,2), and IMEX(4,4,3) in the notation of [5]. In addition, we consider ESDIRK ARK3(2)4L[2]SA from [33], which we abbreviate ARK(4,3), and which demonstrates stability for sufficiently small h .

Figure 8 compares IMEX(1,1,1), IMEX(2,2,2), IMEX(4,4,3) [5], and ARK(4,3) [33] with FIMEX-Radau*(q, κ) for $q \in \{2, 3, 4\}$ and $\kappa \in \{0, 1, 2\}$. Plots present solution accuracy with respect to timestep size, h , and total wallclock time to solution with respect to accuracy, for two different diffusion coefficients. FIMEX-Radau methods are the most efficient in terms of error as a function of runtime to solution for accuracy $\lesssim 10^{-3}$. Moreover, for $\epsilon = 0.1$ and $\epsilon = 10$, the FIMEX-Radau* methods obtain $\sim 10^{-5}$ accuracy, respectively, $\approx 16 \times$ and $\approx 16 \times$ faster than the best IMEX-RK method, IMEX(4,4,3). In terms of order of convergence, we see that all methods suffer from some order-reduction, particularly for $\epsilon = 10$, but FIMEX-Radau* methods retain much smaller error constants than the IMEX-RK methods.

In terms of wallclock time, each iterator application tends to decrease error roughly proportionally with its computational cost. However, they also provide improved stability, e.g., for $\epsilon = 0.1$ and FIMEX-Radau*(4, κ), the method largely diverges for $\kappa = 0$ (except at the smallest h considered, $h = 0.005$), but applying one

FIG. 8. ℓ^2 -accuracy as a function of h and wallclock time to solution.TABLE 1
 ℓ^2 -error and runtime for all stable integrators with $\epsilon = 0.01$ and $h = 0.01$.

	IMEX(1,1,1)	Radau*(2,0)	Radau(2,1)
ℓ^2 -error	0.018	0.009	0.009
Runtime (seconds)	35	37	71
	IMEX(4,4,3)	Radau(3,1)	Radau*(3,2)
ℓ^2 -error	$7.4 \cdot 10^{-5}$	$2.1 \cdot 10^{-5}$	$2.6 \cdot 10^{-6}$
Runtime (seconds)	165	122	151
	Radau*(4,2)		

or two applications of the iterator yields a stable, high-order method. Now consider $\epsilon = 0.01$ and $h = 0.01$. Results for the *only* IMEX-RK methods that did not diverge are shown in Table 1, along with a selection of FIMEX-Radau and FIMEX-Radau* methods that converge. Note that the one exception is FIMEX-Radau*(2,0), which again provides improved accuracy over IMEX-Euler, at a relatively trivial cost, and FIMEX-Radau(2,1), which improves the accuracy of IMEX-Euler via the application of an iterator. Note that the FIMEX-Radau* methods again provide several orders of magnitude smaller error than IMEX(4,4,3) for a comparable wallclock time.

6.3.2. Advection-diffusion-reaction. Here we consider a more practical splitting of explicit treatment for a nonlinear reaction with $\gamma = 10$ and implicit treatment for the advection-diffusion equation, with $\epsilon = 0.1$. By treating the reaction explicitly, we do not have to rebuild the implicit AMG solver at every timestep or nonlinear iteration, which for sparse algebraic solvers is often a nontrivial expense. We also compare

with classical A-stable and L-stable implicit DIRK methods, which use a simplified Newton iteration linearized about the beginning of each timestep (we found this to be most efficient on average, compared with a full Newton iteration, which requires rebuilding the solver every iteration, or a Picard iteration that lags the nonlinearity). DIRK methods are prefixed with the stability (A or L) and suffixed with the order, e.g., LDIRK3 means a third-order L-stable DIRK method. The methods are used as implemented in the MFEM library [2]. The forcing function is chosen for an exact solution $u_*(x, y, t) = 1.0/(e^{10(x+y-t)} + 1)$ for $x, y, t \in [0, 1]$.

For a given time integration scheme, we solve the implicit equations to tolerance 10^{-3} times the order of expected spatial and temporal accuracy based on h_x and h to the appropriate power. This is to ensure systems are solved sufficiently accurately to achieve discretization error but avoid oversolving (e.g., achieving 10 digits of residual accuracy for 2 digits of physical accuracy). Note that here we are using left GMRES acceleration and the preconditioned residual as a stopping tolerance; because the FIRK preconditioners developed in [45, 44] are proven to result in preconditioned operators with condition number bounded $\mathcal{O}(1)$, we expect the preconditioned residual to provide a very good approximation of the error (of course, with an exact preconditioner the preconditioned residual equals the error); similarly for DIRK methods, AMG is a very effective preconditioner for this problem. All simulations are run with 64 total MPI processes. The implicit simplified Newton iteration typically requires 4–10 iterations to converge, where each iteration inverts the (simplified) Jacobian to 3 digits relative residual tolerance.

Error as a function of h and total wallclock time are shown in Figure 9. Here we see that most schemes achieve their expected accuracy (recall the order of $\text{Radau}^*(q, \kappa) = q + \kappa$, up to the order of Radau, given by $2q - 3$). $\text{Radau}^*(q, 0)$ is less predictable (e.g., see $q = 4, 5$ for $\kappa = 0$), but the error constants remain small once a reasonably small stepsize is chosen. It is possible the order-reduction is due to time-dependent

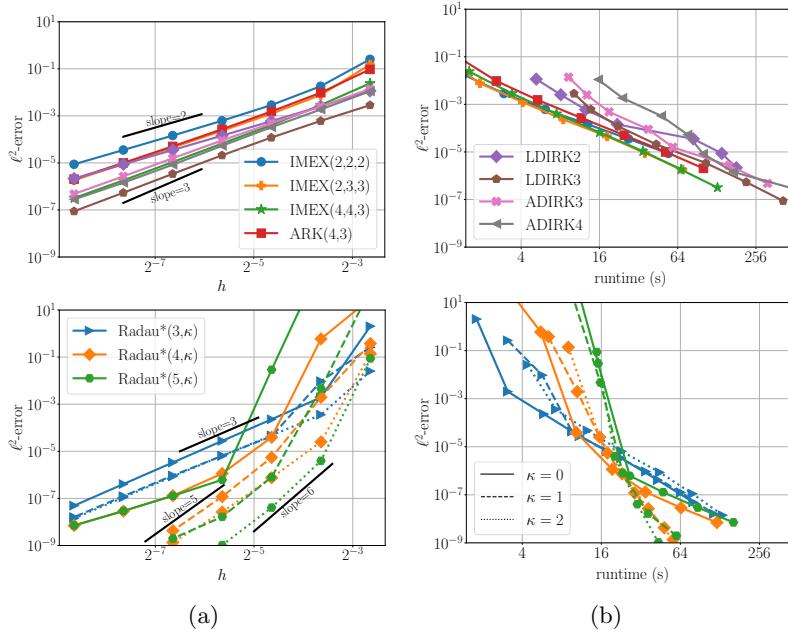


FIG. 9. ℓ^2 -accuracy as a function of h and wallclock time to solution. Legends apply to both images in a row.

boundary conditions, a known problem with RK methods, as we can also see that IMEX-RK and DIRK methods don't quite reach expected third-order convergence. In general, we see that IMEX schemes are advantageous for this problem over DIRK methods, typically achieving better accuracy for a fixed wallclock time. The difference is small for classical IMEX-RK methods but more noticeable with FIMEX-Radau* methods. For moderate to high accuracies, roughly lower than 10^{-4} , the FIMEX-Radau* methods are the best solution with respect to wallclock time. It is also worth pointing out that the asymptotic error constants for FIMEX-Radau* are quite small, e.g., for third-order methods with a fixed h , FIMEX-Radau* error is $10-100\times$ smaller than comparable-order IMEX-RK schemes.

7. Summary and conclusion. In the first half of this paper we introduced a general framework for constructing multivalued or one-step additive polynomial time integrators with any degree of implicitness. By utilizing interpolating polynomials, the framework enables the derivation of high-order methods without requiring the solution of nonlinear order conditions (even if the underlying equation has an arbitrary number of additive partitions). Moreover, by selecting high-order ODE polynomials, one trivially achieves high stage order that prevents order-reduction on stiff problems. Last, the framework also includes iterators that can be used to compute initial conditions and to construct composite PBMs with improved stability and accuracy properties.

The second-half of this paper focused exclusively on FIMEX methods that utilize Adams ODE polynomials. Specifically we introduced a new class of FIMEX-Radau integrators, which use a Radau IIA fully implicit component coupled with a parallel block method for the explicit component. In section 6 we demonstrated the significant potential of these integrators in practice. In each of our experiments, FIMEX-Radau methods consistently yielded the most accurate solution compared with other methods from [5, 33]. Moreover, thanks to recent developments in solvers for fully implicit RK methods [45, 44], they were also the fastest methods (in terms of total runtime) to obtain a given accuracy. Furthermore, by using iterators, the composite FIMEX-Radau methods offer a simple way to develop very high-order FIMEX integrators with low storage cost. We believe these methods are most applicable for problems where moderate to high accuracy is desired, and one can benefit from the strong accuracy and stability provided by combining multistep methods and fully implicit methods (e.g., for time-dependent boundary conditions, or semistiff nonlinearities that one does not want to treat implicitly).

Thus far, we have only introduced one class of additive polynomial integrators. In part two of this work we will introduce new families of diagonally implicit IMEX-polynomial integrators that are based on the BDF ODE polynomials. In addition, we point out that Radau and other fully implicit RK methods offer a number of unique advantages and properties that cannot be obtained with other classes of integrators. Future work will study the potential of FIMEX-Radau and related integrators on topics such as differential algebraic equations and conservation of invariants which fully implicit RK methods are uniquely suited for. We also plan to develop strategies for adaptively controlling the stepsize and the number of iterations κ for polynomial FIMEX collocation methods.

REFERENCES

[1] G. AKRIVIS, O. KARAKASHIAN, AND F. KARAKATSANI, *Linearly implicit methods for nonlinear evolution equations*, Numer. Math., 94 (2003), pp. 403–418.

- [2] R. ANDERSON, J. ANDREJ, A. BARKER, J. BRAMWELL, J.-S. CAMIER, J. CERVENY, V. DOBREV, Y. DUDOUIT, A. FISHER, T. KOLEV, W. PAZNER, M. STOWELL, V. TOMOV, J. DAHM, D. MEDINA, AND S. ZAMPINI, *MFEM: A modular finite element methods library*, Comput. Math. Appl., 81 (2021), pp. 42–74, <https://doi.org/10.1016/j.camwa.2020.06.009>.
- [3] D. N. ARNOLD, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal., 19 (1982), pp. 742–760, <https://doi.org/10.1137/0719052>.
- [4] D. N. ARNOLD, F. BREZZI, B. COCKBURN, AND L. D. MARINI, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2002), pp. 1749–1779, <https://doi.org/10.1137/S0036142901384162>.
- [5] U. M. ASCHER, S. J. RUUTH, AND R. J. SPITERI, *Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations*, Appl. Numer. Math., 25 (1997), pp. 151–167.
- [6] U. M. ASCHER, S. J. RUUTH, AND B. T. WETTON, *Implicit-explicit methods for time-dependent partial differential equations*, SIAM J. Numer. Anal., 32 (1995), pp. 797–823.
- [7] S. BOSCARINO, *Error analysis of IMEX Runge-Kutta methods derived from differential-algebraic systems*, SIAM J. Numer. Anal., 45 (2007), pp. 1600–1621.
- [8] J. C. BUTCHER AND N. GOODWIN, *Numerical Methods for Ordinary Differential Equations*, Vol. 2, Wiley, New York, 2016.
- [9] T. BUVOLI, *Polynomial-Based Methods for Time-Integration*, Ph.D. thesis, University of Washigton, 2018.
- [10] T. BUVOLI, *A class of exponential integrators based on spectral deferred correction*, SIAM J. Sci. Comput., 42 (2020), pp. A1–A27.
- [11] T. BUVOLI, *Constructing Polynomial Block Methods*, preprint, arXiv:2011.00671, 2020.
- [12] T. BUVOLI, *Codebase for “Additive Polynomial Time Integrators, Part I: Framework and Fully-Implicit-Explicit (FIMEX) Collocation Methods”*, <https://github.com/buvoli/apbm>, 2021.
- [13] T. BUVOLI, *Exponential polynomial block methods*, SIAM J. Sci. Comput., 43 (2021), pp. A1692–A1722.
- [14] T. BUVOLI AND M. TOKMAN, *Constructing new time integrators using interpolating polynomials*, SIAM J. Sci. Comput., 41 (2019), pp. A2911–A2937.
- [15] M. CALVO, J. DE FRUTOS, AND J. NOVO, *Linearly implicit Runge-Kutta methods for advection-reaction-diffusion equations*, Appl. Numer. Math., 37 (2001), pp. 535–549.
- [16] A. CARDONE, Z. JACKIEWICZ, A. SANDU, AND H. ZHANG, *Extrapolation-based implicit-explicit general linear methods*, Numer. Algorithms, 65 (2014), pp. 377–399.
- [17] A. CARDONE, Z. JACKIEWICZ, A. SANDU, AND H. ZHANG, *Construction of highly stable implicit-explicit general linear methods*, in Conference Publications, American Institute of Mathematical Sciences, Pasadena, CA, 2015, pp. 185–194.
- [18] H. CHEN, *A splitting preconditioner for the iterative solution of implicit Runge-Kutta and boundary value methods*, BIT, 54 (2014), pp. 607–621.
- [19] A. J. CHRISTLIEB, C. B. MACDONALD, AND B. W. ONG, *Parallel high-order integrators*, SIAM J. Sci. Comput., 32 (2010), pp. 818–835.
- [20] B. COCKBURN AND C.-W. SHU, *Runge-Kutta discontinuous Galerkin methods for convection-dominated problems*, J. Sci. Comput., 16 (2001), pp. 173–261, <https://doi.org/10.1023/a:1012873910884>.
- [21] E. M. CONSTANTINESCU AND A. SANDU, *Extrapolated implicit-explicit time stepping*, SIAM J. Sci. Comput., 31 (2010), pp. 4452–4477.
- [22] G. COOPER AND A. SAYFY, *Additive methods for the numerical solution of ordinary differential equations*, Math. Comp., 35 (1980), pp. 1159–1172.
- [23] G. DIMARCO AND L. PARESCHI, *Implicit-explicit linear multistep methods for stiff kinetic equations*, SIAM J. Numer. Anal., 55 (2017), pp. 664–690.
- [24] A. DUTT, L. GREENGARD, AND V. ROKHLIN, *Spectral deferred correction methods for ordinary differential equations*, BIT, 40 (2000), pp. 241–266.
- [25] R. D. FALGOUT AND U. M. YANG, *hypre: A library of high performance preconditioners*, in Proceedings of the European Conference on Parallel Processing, Lecture Notes in Comput. Sci. 2331, 2002, pp. 632–641.
- [26] P. E. FARRELL, R. C. KIRBY, AND J. MARCHENA-MENENDEZ, *Irksome: Automating Runge-Kutta Time-Stepping for Finite Element Methods*, preprint, arXiv:2006.16282, 2020.
- [27] R. GLANDON, M. NARAYANAMURTHI, AND A. SANDU, *Linearly Implicit Multistep Methods for Time Integration*, preprint, arXiv:2011.10685, 2020.
- [28] E. HAIRER, S. P. NØRSETT, AND G. WANNER, *Solving Ordinary Differential Equations I: Nonsstiff Problems*, Springer Series in Computational Mathematics, Springer Berlin, Heidelberg, 1993, 528.

- [29] E. HAIRER AND G. WANNER, *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*, Springer Series in Computational Mathematics, Springer Berlin, Heidelberg, 1996, 614.
- [30] E. HAIRER AND G. WANNER, *Stiff differential equations solved by Radau methods*, J. Comput. Appl. Math., 111 (1999), pp. 93–111.
- [31] G. IZZO AND Z. JACKIEWICZ, *Highly stable implicit-explicit Runge-Kutta methods*, Appl. Numer. Math., 113 (2017), pp. 71–92.
- [32] X. JIAO, X. WANG, AND Q. CHEN, *Optimal and Low-Memory Near-Optimal Preconditioning of Fully Implicit Runge-Kutta Schemes for Parabolic PDEs*, preprint, arXiv:2012.12779, 2020.
- [33] C. A. KENNEDY AND M. H. CARPENTER, *Additive Runge-Kutta schemes for convection-diffusion-reaction equations*, Appl. Numer. Math., 44 (2003), pp. 139–181.
- [34] C. A. KENNEDY AND M. H. CARPENTER, *Higher-order additive Runge-Kutta schemes for ordinary differential equations*, Appl. Numer. Math., 136 (2019), pp. 183–205.
- [35] A. T. LAYTON AND M. MINION, *Implications of the choice of quadrature nodes for Picard integral deferred corrections methods for ordinary differential equations*, BIT, 45 (2005), pp. 341–373.
- [36] T. A. MANTEUFFEL, J. W. RUGE, AND B. S. SOUTHWORTH, *Nonsymmetric algebraic multigrid based on local approximate ideal restriction (ℓ AIR)*, SIAM J. Sci. Comput., 40 (2018), pp. A4105–A4130.
- [37] M. MINION, *Semi-implicit spectral deferred correction methods for ordinary differential equations*, Commun. Math. Sci., 1 (2003), pp. 471–500.
- [38] W. PAZNER AND P.-O. PERSSON, *Stage-parallel fully implicit Runge-Kutta solvers for discontinuous Galerkin fluid simulations*, J. Comput. Phys., 335 (2017), pp. 700–717.
- [39] M. M. RANA, V. E. HOWLE, K. LONG, A. MEEK, AND W. MILESTONE, *A New Block Preconditioner for Implicit Runge-Kutta Methods for Parabolic PDE*, preprint, arXiv:2010.11377, 2020.
- [40] A. SANDU AND M. GÜNTHER, *A generalized-structure approach to additive Runge-Kutta methods*, SIAM J. Numer. Anal., 53 (2015), pp. 17–42.
- [41] L. F. SHAMPINE AND H. WATTS, *Block implicit one-step methods*, Math. Comp., 23 (1969), pp. 731–740.
- [42] A. A. SIVAS, B. S. SOUTHWORTH, AND S. RHEBERGEN, *AIR algebraic multigrid for a space-time hybridizable discontinuous galerkin discretization of advection(-diffusion)*, SIAM J. Sci. Comput., 43 (2021), pp. A3393–A3416.
- [43] B. SOLEIMANI AND R. WEINER, *Superconvergent IMEX peer methods*, Appl. Numer. Math., 130 (2018), pp. 70–85.
- [44] B. S. SOUTHWORTH, O. A. KRZYSIK, AND W. PAZNER, *Fast Solution of Fully Implicit Runge-Kutta and Discontinuous Galerkin in Time for Numerical PDEs, Part II: Nonlinearities and DAEs*, preprint, arXiv:2101.01776, 2021.
- [45] B. S. SOUTHWORTH, O. A. KRZYSIK, W. PAZNER, AND H. D. STERCK, *Fast Solution of Fully Implicit Runge-Kutta and Discontinuous Galerkin in Time for Numerical PDEs, Part I: The Linear Setting*, preprint, arXiv:2101.00512, 2021.
- [46] D. WANG AND S. J. RUUTH, *Variable step-size implicit-explicit linear multistep methods for time-dependent partial differential equations*, J. Comput. Math., 26 (2008), pp. 838–855.
- [47] G. WANNER, *A short proof on nonlinear A-stability*, BIT, 16 (1976), pp. 226–227.
- [48] H. ZHANG, A. SANDU, AND S. BLAISE, *High order implicit-explicit general linear methods with optimized stability regions*, SIAM J. Sci. Comput., 38 (2016), pp. A1430–A1453.