AIAA SciTech Forum
23-27 January 2023, National Harbor, MD & Online
AIAA SCITECH 2023 Forum

10.2514/6.2023-0662

Pseudo-time Stepping Strategies for Space-Time Discontinuous Galerkin Discretizations

Sung-Hwan Yoon[1], Andrew C. Kirby[2] and Dimitri J. Mavriplis[3]
Department of Mechanical Engineering, University of Wyoming, Laramie, Wyoming, 82071-3295

Abstract

In this study, we present strategies to improve the computational efficiency of high-order space-time DG. The relation between DG in time and implicit Runge-Kutta methods is discussed further and efficient smoothing strategies for solving the implicit systems at each time step are investigated. We suggest a simple IMEX multi-stage optimized RK scheme for pseudo-time stepping that treats the spatial residual explicitly and the temporal residual implicitly. An IMEX four-stage optimized RK scheme has been tested for the isentropic vortex problem at high spatial and temporal orders. For small physical time steps, the IMEX scheme provides an 8.7x speedup over the purely explicit pseudo-time stepping scheme. For large physical time steps, the IMEX scheme reduces to an explicit RK scheme optimized for stability in pseudo-time and provides 7.8x faster convergence rate than single-stage RK scheme in pseudo-time. The performance of the IMEX pseudo-time stepping scheme is examined for flow over a NACA0012 airfoil in 2D and 3D as a function of the physical time step size and variations in grid resolution, showing superior performance to the equivalent explicit RK scheme using either local or global minimum time stepping in pseudo time.

I. Introduction

The increasing interest in turbulence scale-resolving methods for computational aerodynamics is largely due to potential of these methods to more accurately predict flows with significant regions of separation. While, to date, many of the proposed scale-resolving methods rely on explicit schemes in time, there are good reasons to consider implicit time-stepping methods for scale-resolving simulations. In particular, for industrial simulations covering a wide range of scales, computational meshes with wide variations in spatial resolution are the norm, and these can hinder the efficiency of explicit time-stepping methods, where the global time step is determined by the smallest cell in the mesh. However, the time-step requirements for scale-resolving simulations can be significantly different than those for unsteady Reynolds-averaged Navier-Stokes (RANS) methods, since one can expect smaller time steps and higher temporal accuracy to be required for capturing the wide range of temporal scales of the resolved eddies in turbulent flows. Thus, an implicit time-stepping approach for scale resolving simulations must be cost competitive with explicit time-stepping schemes, not only for large physical time steps, but also in the limit of small physical time steps, and these must remain stable and efficient on meshes of widely varying spatial resolution.

Space-time discontinuous Galerkin (DG) methods are an extension into the time dimension of high-order DG methods in space which have been widely studied [1, 2] and proposed as a suitable approach for time-implicit scale-resolving methods. Space-time DG methods share the attractive properties of DG in space and add more flexibility in the temporal dimension in terms of accuracy (p-order) and the adaptivity of p-order. A weakness of high-order space-time DG is expensive computational and storage costs. As the p-order increases, the additional degrees of freedom in the time dimension quickly make the method impractical in many engineering applications, especially

[1] Assistant Research Scientist; email: sung.hwan.yoon@uwyo.edu
[2] Research Scientist, AIAA member; email: akirby@uwyo.edu
[3] Professor, AIAA Associate Fellow; email: mavripl@uwyo.edu

three-dimensional spatial problems. Recent studies have been trying to overcome the weakness by using tensor-product formulations: tensor-product based sum factorizations and preconditioners [3–5].

In previous work [6], we compared DG in time (DGT) discretizations with FIRK (fully implicit Runge-Kutta) schemes and showed that DGT takes an analogous form to a preconditioned FIRK scheme or *p*FIRK. It was also shown that these temporal discretizations can be solved at each time step using an explicit pseudo-time stepping approach, which can be accelerated using a *p*-multigrid solver in pseudo-time

The present work is a continuation of our efforts to improve the computational efficiency of high-order space-time DG. The relation between DGT and FIRK is discussed further, examining the accuracy and stability properties of DGT with different quadrature points. Efficient smoothing strategies in pseudo-time are developed for solving the implicit system at each time step for these discretizations at various orders of accuracy. These are used to accelerate the solution of space-time systems at high spatial and temporal orders of accuracy. Techniques for further accelerating convergence using the proposed pseudo-time stepping methods as smoothers for a *p*-multigrid solver are also discussed.

II. Space-Time Discontinuous Galerkin Discretization Using Tensor-Product Formulation

II-A. Semi-Discrete Spatial DG Discretization

Let us consider the multi-dimensional hyperbolic conservation laws as follows.

$$\frac{\partial \mathbf{Q}(\mathbf{x},t)}{\partial t} + \nabla \cdot \mathbf{F}(\mathbf{Q}) = 0, \tag{1}$$

where $\mathbf{Q}$ is the state variable vector and $\mathbf{F}$ is the flux function vector. When we multiply by a set of test functions $\phi(\mathbf{x})$ and integrate over a space element, the weak statement for the governing equation can be expressed as:

$$\int_{\Omega_E} \left[\frac{\partial \mathbf{Q}(\mathbf{x},t)}{\partial t} + \nabla \cdot \mathbf{F}(\mathbf{Q})\right] \phi(\mathbf{x})\, d\Omega_E = 0. \tag{2}$$

Integrating by parts and applying Green's theorem yields:

$$\int_{\Omega_E} \frac{\partial \mathbf{Q}(\mathbf{x},t)}{\partial t} \phi(\mathbf{x}) d\Omega_E - \int_{\Omega_E} (\mathbf{F} \cdot \nabla)\phi(\mathbf{x}) d\Omega_E + \int_{\partial\Omega_E} (\mathbf{F} \cdot \mathbf{n})\, \phi(\mathbf{x})\, d(\partial\Omega_E) = 0. \tag{3}$$

The state vector is written as an expansion in terms of the basis functions and the solution degrees of freedom in space as:

$$\mathbf{Q} = \sum_{ij} \mathbf{Q}_{ij} \phi_i(x) \phi_j(y)$$

The first term is an integral of the temporal derivative, and it can be simplified as

$$\int_{\Omega_E} \frac{\partial \mathbf{Q}(\mathbf{x},t)}{\partial t} \phi(\mathbf{x}) d\Omega_E = \mathbf{M} \frac{\partial \mathbf{Q}_{ij}}{\partial t} \tag{4}$$

where $\mathbf{M}$ is the mass matrix defined as $\mathbf{M} = \int \phi_i\, \phi_j d\Omega_E$ for two-dimensional space. The second and third term of Eq.(3) are volume and surface integrals, respectively. We can define a spatial residual, $\mathbf{R}_{ij}(\mathbf{Q})$, into which the volume and the surface integrals are combined as follows.

$$\mathbf{R}_{ij}(\mathbf{Q}) = -\int_{\Omega_E} (\mathbf{F} \cdot \nabla)\phi(\mathbf{x}) d\Omega_E + \int_{\partial\Omega_E} (\mathbf{F} \cdot \mathbf{n})\, \phi(\mathbf{x})\, d(\partial\Omega_E) \tag{5}$$

Then, we can invoke isoparametric mapping and expand the solution in order to obtain a system of algebraic equations to be solved. The details of the expansion and integration processes for the tensor-product form of basis

functions follow a standard approach and the specifics of our implementation are explained in [7]. Substituting Eqs. (4) and (5) into Eq. (3) and omitting the ij-index for simplicity, a semi-discrete formulation can be written as:

$$\mathbf{M}\frac{\partial \mathbf{Q}}{\partial t} + \mathbf{R}(\mathbf{Q}) = 0 \tag{6}$$

In the present paper, we focus on formulations in two-dimensional space for simplicity. Nevertheless, the derived formulations can be extended to three-dimensional space in a straight-forward manner.

II-B. Unsteady Residual for Space-Time DG Formulation

In order to introduce the differences in the formulations between space DG and space-time DG, rather than a compact way described in [8], we derive the space-time DG formulation by multiplying the space DG formulation given as Eq. (3) by an additional basis function for the time dimension $\psi(t)$ and integrating it over a time interval $I^n = [t^n, t^{n+1}]$, where $t^{n+1} = t^n + \Delta t$.

$$\int_{I^n}\left[\int_{\Omega_E}\frac{\partial \mathbf{Q}(\mathbf{x},t)}{\partial t}\phi(\mathbf{x})d\Omega_E - \int_{\Omega_E}(\mathbf{F}\cdot\nabla)\phi(\mathbf{x})d\Omega_E + \int_{\partial\Omega_E}(\mathbf{F}\cdot\mathbf{n})\,\phi(\mathbf{x})\,d(\partial\Omega_E)\right]\psi(t)dt = 0 \tag{7}$$

Applying Eq.(5) yields

$$\int_{I^n}\int_{\Omega_E}\frac{\partial \mathbf{Q}(\mathbf{x},t)}{\partial t}\phi(\mathbf{x})\psi(t)d\Omega_E dt + \int_{I^n}\mathbf{R}(\mathbf{Q})\psi(t)dt = 0. \tag{8}$$

The first term can be written as:

$$\int_{I^n}\int_{\Omega_E}\frac{\partial \mathbf{Q}(\mathbf{x},t)}{\partial t}\phi(\mathbf{x},t)d\Omega_E dt = \int_{I^n}\int_{\Omega_E}\frac{\partial(\mathbf{Q}\phi)}{\partial t}d\Omega_E dt - \int_{I^n}\int_{\Omega_E}\mathbf{Q}\frac{\partial\phi}{\partial t}d\Omega_E dt \tag{9}$$

$$= \int_{\Omega_E}[\mathbf{Q}(\mathbf{x},t_-^{n+1})\phi(t_-^{n+1}) - \mathbf{Q}(\mathbf{x},t_-^{n})\phi(t_+^{n})]\,d\Omega_E - \int_{I^n}\int_{\Omega_E}\mathbf{Q}\frac{\partial\phi}{\partial t}d\Omega_E dt \tag{10}$$

where $\phi(\mathbf{x},t) = \phi(\mathbf{x})\psi(t)$. Previously in [6], the temporal basis function was assumed to be of the same form as the spatial basis function: $\psi(t) = \phi(t)$ with Gauss-Legendre quadrature points. However, it is found that the characteristics of the temporal DG scheme (DGT) may vary according to the choice of temporal basis function (quadrature points, more specifically). Here, we keep the general form of the temporal basis function $\psi(t)$. Details will be discussed with the relation between DGT and FIRK in the next sub-section.

Herein, let us define the left-hand-side of the Eq.(8) as an unsteady residual that includes the temporal discretization as well as spatial discretization. Considering the semi-discrete formulation Eq.(6), the additional integration over a time interval I^n needs to be normalized by dividing Δt for a consistent definition of the unsteady residual. Then, the unsteady residual $\mathbf{L}(\mathbf{Q})$ for space-time DG can be defined as follows.

$$\text{DG in time:}\quad \mathbf{L}(\mathbf{Q}^{n+1}) = \frac{1}{\Delta t}\left[\int_{\Omega_E}[\mathbf{Q}(\mathbf{x},t_-^{n+1})\psi(t_-^{n+1}) - \mathbf{Q}(\mathbf{x},t_-^{n})\psi(t_+^{n})]\,d\Omega_E - \int_{I^n}\int_{\Omega_E}\mathbf{Q}^{n+1}\frac{\partial\psi}{\partial t}d\Omega_E dt\right] + \frac{1}{\Delta t}\int_{I^n}\mathbf{R}(\mathbf{Q}^{n+1})\psi(t)dt \tag{11}$$

We define a sub-vector that consists of quadrature point values for spatial DG as $\mathbf{q} = Q_{ij}$, then the quadrature point values in a space-time element can also be expressed as $\mathbf{q}_k = Q_{ijk}$. When this is expanded out, a matrix form of $\mathbf{L}(\mathbf{Q})$ for the case of $p_t = 1$ (where p_t corresponds to the temporal polynomial order) is given as:

$$\text{DG in time:} \quad \mathbf{L}(\mathbf{Q}^{n+1}) = \frac{\mathbf{M}}{\Delta t}\left\{(\boldsymbol{\Phi}^{n+1} - \boldsymbol{D}_{\boldsymbol{\omega}})\begin{bmatrix}\mathbf{q}_1^{n+1}\\ \mathbf{q}_2^{n+1}\end{bmatrix} - \widehat{\boldsymbol{\Psi}}^{n}\begin{bmatrix}\widehat{\mathbf{q}}^{n}\\ \widehat{\mathbf{q}}^{n}\end{bmatrix}\right\} + \frac{\boldsymbol{\Omega}}{2}\begin{bmatrix}\mathbf{r}_1^{n+1}\\ \mathbf{r}_2^{n+1}\end{bmatrix} \tag{12}$$

where $\boldsymbol{\Phi}^{n+1}$, $\boldsymbol{D}_{\boldsymbol{\omega}}$, $\widehat{\boldsymbol{\Psi}}^{n}$ and $\boldsymbol{\Omega}$ are matrices calculated using temporal basis function $\psi(t)$. The definitions and details of derivation process are given in [6].

II-C. Relation Between DGT and FIRK

The general form of an implicit Runge-Kutta scheme for solving the equation

$$\frac{dq}{dt} = f(t, q)$$

is given as

$$q_i = q^n + \Delta t \sum_{j=1}^{s} a_{ij} f\left(t^n + c_j \Delta t, q_j\right) \quad i = 1,2, \dots s$$

$$q^{n+1} = q^n + \Delta t \sum_{i=1}^{s} b_i f(t^n + c_i \Delta t, q_i)$$

Here the first equation represents the calculation of the q_i stage values, while the second equation computes the value q^{n+1} at the end of the time interval given the stage values. The coefficients of the scheme are defined in the Butcher tableau as

$$\begin{array}{c|c} \boldsymbol{c} & \boldsymbol{A} \\ \hline & \boldsymbol{b^T} \end{array} = \begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1s} \\ c_2 & a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & a_{s2} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array}$$

Using an IRK scheme to compute the time evolution of Eq.(6) results in a similar space-time residual as the DGT scheme, corresponding the equations which must be solved to obtain the IRK stage values. For a 2-stage IRK scheme, these equations can be written as:

$$\mathbf{L}(\mathbf{Q}^{n+1}) = \frac{\mathbf{M}}{\Delta t}\left\{\begin{bmatrix}\mathbf{q}_1^{n+1}\\ \mathbf{q}_2^{n+1}\end{bmatrix} - \begin{bmatrix}\widehat{\mathbf{q}}^{n}\\ \widehat{\mathbf{q}}^{n}\end{bmatrix}\right\} + \boldsymbol{A}\begin{bmatrix}\mathbf{r}_1^{n+1}\\ \mathbf{r}_2^{n+1}\end{bmatrix} \tag{13}$$

where the entries of the matrix $\boldsymbol{A}$ correspond to the a_{ij} coefficients in the Butcher tableau of the IRK scheme. In reference [9] and [6], a preconditioned form of the IRK scheme was considered, which is obtained by multiplying through by $\boldsymbol{A}^{-1}$, resulting in the preconditioned space-time residual:

$$\mathbf{L}(\mathbf{Q}^{n+1}) = \frac{\mathbf{M}}{\Delta t}\left\{\boldsymbol{A}^{-1}\begin{bmatrix}\mathbf{q}_1^{n+1}\\ \mathbf{q}_2^{n+1}\end{bmatrix} - \boldsymbol{A}^{-1}\begin{bmatrix}\widehat{\mathbf{q}}^{n}\\ \widehat{\mathbf{q}}^{n}\end{bmatrix}\right\} + \begin{bmatrix}\mathbf{r}_1^{n+1}\\ \mathbf{r}_2^{n+1}\end{bmatrix} \tag{14}$$

The preconditioned form of the IRK scheme given above has obvious similarities to the DG-in-time formulation given in equation (12). In reference [10], a proof is given which shows that a discontinuous Galerkin-in-time discretization using right Radau quadrature rules corresponds to the well-known Radau IIA implicit Runge-Kutta scheme. In fact, it is relatively simple to show that equation (12) above reduces to the above expression for a 2-stage Radau IIA IRK scheme when the ψ(t) basis functions are taken as the right Radau polynomials and the quadrature is performed using the corresponding right Radau points. Reference [19] extends the proofs of reference [10] to further derive a complete class of IRK-DG schemes based on the particular quadrature rules used in the DG scheme. Using Gauss quadrature in time, which matches the current quadrature used in the spatial dimensions, produces a so-called

DG-Gauss IRK method, which is different from the Gauss IRK method found in the literature. From the IRK literature [8, 11-12], for non-stiff problems s-stage Gauss IRK schemes are known to deliver $2s=2p_t+2$ accuracy and to be A-stable but not L-stable. On the other hand, the Radau IIA schemes achieve one order of accuracy lower than their Gauss counterparts, i.e. $2s-1=2p_t+1$, but are both A-stable and L-stable. In reference [19] it is shown that all DG-IRK schemes are both A-stable and L-stable and achieve $2s-1=2p_t+1$ accuracy, similar to the Radau IIA IRK scheme.

Numerical tests performed in this work have confirmed the expected design accuracy of the various IRK schemes and DG-in-time schemes for the canonical test case of an isentropic vortex convection problem [13–15], which is shown in section IV in detail. As discussed therein, our preferred temporal scheme remains the IRK Radau IIA scheme, due to its favorable stability and stiff-order reduction properties [18].

III. Pseudo-Time Stepping

In the above section, we only considered the accuracy and stability of the various temporal schemes and assumed the implicit system arising at each time step was solved to machine precision. In the following sections we focus on the techniques for solving the implicit systems generated by these discretizations at each time step.

Once we have formulated $\mathbf{L}(\mathbf{Q}^{n+1})$ for any space-time discretization, we are interested in solving $\mathbf{L}(\mathbf{Q}^{n+1}) = 0$ at every time step to advance in time. In order to solve this non-linear set of equations, pseudo-time stepping, also known as dual-time stepping, is an often used approach which seeks to mimic the physical time evolution of the solution by adding a pseudo-time term $\mathbf{M}\frac{\partial \mathbf{Q}}{\partial \tau}$ as [16]:

$$\mathbf{M}\frac{\partial \mathbf{Q}}{\partial \tau} + \mathbf{L}(\mathbf{Q}^{n+1}) = 0 \tag{15}$$

and integrating in pseudo-time τ until steady-state in pseudo time is achieved. This nonlinear system can be solved in pseudo-time using simple explicit time-stepping methods, which in turn may be accelerated using multigrid methods. A simple approach is the explicit single-stage RK (or the 1st order accurate forward Euler scheme) time stepping in pseudo-time, which gives

$$\frac{\mathbf{M}}{\Delta\tau}(\mathbf{Q}^{k+1} - \mathbf{Q}^{k}) + \mathbf{L}(\mathbf{Q}^{n+1}) = 0 \tag{16}$$

Here, $\mathbf{Q}^k$ is the pseudo-time level and $\mathbf{Q}^n$ is the physical-time level such that

$$\mathbf{Q}^{k+1} \approx \mathbf{Q}^{k} \rightarrow \mathbf{Q}^{n+1} \text{ as } k \rightarrow \infty \text{ and } \mathbf{L}(\mathbf{Q}^{k}) \rightarrow 0 \tag{17}$$

Thus, we can write Eq. (16) as:

$$\mathbf{Q}^{k+1} = \mathbf{Q}^{k} - \Delta\tau\mathbf{M}^{-1}\mathbf{L}(\mathbf{Q}^{k}) \tag{18}$$

This gives an explicit equation to be solved for $\mathbf{Q}^{k+1}$.

III-A. Explicit Multi-Stage Optimized RK in Pseudo-Time

The explicit single-stage RK scheme can be easily extended to explicit multi-stage schemes in pseudo-time. A general *s*-stage explicit RK scheme can be represented by the Butcher tableau [13]. Vermeire et al. [17] optimized RK schemes for pseudo-time stepping for high-order spatial discretizations. Since the temporal accuracy of pseudo-time stepping is not a consideration, they maximized the stability region while trading off the accuracy as 1st order accurate. The methods are optimal in that they allow the largest possible pseudo time step $\Delta\tau$ to be taken for a given number of RK stages. As a result, they observed speedup factors of 1.8x for a flux-reconstruction DG spatial discretization with their optimized 4-stage RK scheme in terms of maximum effective time step size.

Butcher tableau coefficients for the optimized RK schemes are provided online as Electronic Supplementary Material with the article in reference [17]. The optimized 4-stage RK scheme for FRDG ($p_s = 5$), for example, has the coefficients as follows:

0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
0.40128709	0.40128709	0.00000000	0.00000000	0.00000000
0.56449983	0.28224991	0.28224991	0.00000000	0.00000000
0.87678807	0.25972925	0.25479937	0.36225945	0.00000000
	0.20334721	0.19932974	0.28339585	0.31392720

The actual performance of the optimized 4-stage RK schemes for numerical tests will be discussed in section IV.

III-B. IMEX Multi-Stage Optimized RK Schemes

In ewference [17], the maximum stable pseudo-time step $\Delta\tau$ was obtained empirically and assumed to be related to the explicit stability limit due to the spatial residual term. However, the unsteady residual $\mathbf{L}(\mathbf{Q})$ consists of temporal and spatial components. In our previous work [6], we considered the effect of the temporal component, which can become more restrictive particularly for small physical time steps.

For illustration purposes, we consider the case where a BDF1 temporal discretization is employed (i.e. corresponding to $p_t = 0$ DG-in-time). In this case, the explicit pseudo-time stepping scheme becomes:

$$\mathbf{Q}^{k+1} = \mathbf{Q}^{\boldsymbol{k}} - \frac{\Delta\tau}{\Delta t}[\mathbf{Q}^k - \mathbf{Q}^{\boldsymbol{n}}] - \Delta\tau\mathbf{M}^{-1}\mathbf{R}(\mathbf{Q}^k) \tag{19}$$

As can be seen from this simple example, for small Δt, the second term of the right-hand side becomes large and may require a smaller pseudo-time step to guarantee stability. An alternative approach which enables the use of a constant pseudo-time step regardless of the size of the physical time step is to treat this term implicitly in pseudo-time as:

$$\left[1 + \frac{\Delta\tau}{\Delta t}\right]\mathbf{Q}^{k+1} = \mathbf{Q}^k + \frac{\Delta\tau}{\Delta t}\mathbf{Q}^n - \Delta\tau\mathbf{M}^{-1}\mathbf{R}(\mathbf{Q}^k) \tag{20}$$

noting that the second term of the right-hand side of Eq.(20) is a constant source term, involving the known values of $\mathbf{Q}$ at the beginning of the time interval. In previous work [6], we extended this to DG-in-time and IRK temporal discretizations as:

$$\left[1 + \frac{\Delta\tau}{\Delta t}\boldsymbol{A}^{-1}\right]\mathbf{Q}^{k+1} = \mathbf{Q}^k + \frac{\Delta\tau}{\Delta t}\boldsymbol{A}^{-1}\mathbf{Q}^n - \Delta\tau\mathbf{M}^{-1}\mathbf{R}(\mathbf{Q}^k) \tag{21}$$

where $\boldsymbol{A}$ represents the matrix of IRK coefficients, and $\mathbf{Q}$ represents the vector of stage values for the corresponding IRK scheme (with the analogous representation for DGT schemes). Although this approach was shown to work well for DG-in-time and IRK temporal schemes up to high order, it was only applied for cases using RK1 (simple forward Euler explicit) pseudo-time stepping. In order to extend this approach for multi-stage explicit pseudo-time stepping, which is necessary for the use of the optimized pseudo-time stepping schemes discussed in the previous section, a more general approach must be devised. In this work, we have used the idea of implicit-explicit (also known as IMEX) schemes [11] for extending the implicit treatment of the temporal component of the unsteady residual $\mathbf{L}(\mathbf{Q})$ to multi-stage optimized RK pseudo-time stepping schemes. We proceed by dividing the unsteady residual into temporal and spatial components denoted as $\mathbf{L}(\mathbf{Q}) = \mathbf{M}\{\mathbf{T}(\mathbf{Q}) + \mathbf{S}(\mathbf{Q})\}$. An IMEX scheme that treats the temporal part implicitly and the spatial part explicitly can then be written as:

$$\text{Stage values:} \quad \boldsymbol{Q}^{k,i} = \boldsymbol{Q}^n - \Delta\tau\sum_{j=1}^{i} a_{ij}\,\boldsymbol{T}(\boldsymbol{Q}^{k,j}) - \Delta\tau\sum_{j=1}^{i-1}\hat{a}_{ij}\,\boldsymbol{S}(\boldsymbol{Q}^{k,j}) \tag{22}$$

$$\text{Final new time value:} \quad \boldsymbol{Q}^{n+1} = \boldsymbol{Q}^n - \Delta\tau\sum_{i=1}^{s} b_i\,\boldsymbol{T}(\boldsymbol{Q}^{k,i}) - \Delta\tau\sum_{i=1}^{s}\hat{b}_i\,\boldsymbol{S}(\boldsymbol{Q}^{k,i}) \tag{23}$$

The $\hat{a}$, $\hat{b}$ coefficients determine the explicit scheme, while the corresponding a, b coefficients determine the implicit scheme. Here, superscript k indicates the pseudo-time level and i and j indicate the stage level of the RK scheme. For example, when BDF1 is used as a physical time discretization, we have

BDF1:
$$\boldsymbol{T}(\mathbf{Q}^{k,i}) = \frac{1}{\Delta t}(\mathbf{Q}^{k,i} - \mathbf{Q}^{n}),$$
$$\boldsymbol{S}(\mathbf{Q}^{k,i}) = \boldsymbol{M^{-1}R}(\mathbf{Q}^{k,i})$$

Similarly, we have the following for the preconditioned IRK Radau2A scheme:

*p*FIRK Radau2A:
$$\boldsymbol{T}(\mathbf{Q}^{k,i}) = \frac{1}{\Delta t}\boldsymbol{A^{-1}}(\mathbf{Q}^{k,i} - \mathbf{Q}^{n}),$$
$$\boldsymbol{S}(\mathbf{Q}^{k,i}) = \boldsymbol{M^{-1}R}(\mathbf{Q}^{k,i})$$

In order to devise our IMEX scheme, we recall that we are not interested in pseudo-time accuracy, but only stability for these optimized schemes. Since we already have an explicit multi-stage scheme optimized for stability which operates on the spatial residual, the task is to find a corresponding multi-stage implicit scheme to build the complete IMEX scheme. The simplest approach is to construct a diagonally implicit RK scheme with a single diagonal coefficient at each stage. For this, we assume that the a_{ij} coefficients are all zero except for the diagonal terms: $a_{ii} = c_i$, as required by the consistency condition for RK schemes [18]. Here, we use the optimized RK41 scheme to define the explicit scheme. Then, we have:

Stage values:
$$[\boldsymbol{Q}^{k,i} + \Delta\tau c_i \boldsymbol{T}(\boldsymbol{Q}^{k,i})] = \boldsymbol{Q}^{n} - \Delta\tau \sum_{j=1}^{i-1} \hat{a}_{ij}\, \boldsymbol{S}(\boldsymbol{Q}^{k,j}) \tag{24}$$

Final new time value:
$$\left[\boldsymbol{Q}^{k+1} + \Delta\boldsymbol{\tau}\boldsymbol{T}\left(\boldsymbol{Q}^{k+1}\right)\right] = \boldsymbol{Q}^{n} - \Delta\boldsymbol{\tau} \sum_{i=1}^{s} \widehat{\boldsymbol{b}}_i\, \boldsymbol{S}\left(\boldsymbol{Q}^{k,i}\right) \tag{25}$$

We note that this formulation corresponds to the application of a BDF1 scheme for obtaining the stage values of the implicit component of the IMEX scheme. Since the explicit component is stable by design, the combination of both schemes in this IMEX scheme is guaranteed to be stable for the calculation of the stage values. For the final new time value, the spatial component contribution is assembled using the $\widehat{b_i}$ coefficients of the underlying explicit scheme, whereas the temporal component is treated implicitly as another stage value. This can be interpreted as an extra stage that produces the value of Q at the end of the time interval. Furthermore, for large physical time step values, the implicit terms become negligible and the explicit multi-stage scheme in pseudo-time is recovered.

In order to solve the left-hand side of Eq. (24) and (25), we need to perform a matrix inversion operation. For practical ranges of $the\ p_t$ order, however, the left-hand side of Eq. (24) and (25) consist of very small size matrices. For example, when we use sub-vector for the spatial part, the matrix to be inverted is $N \times N$ where $N = p_t + 1$. Up to $p_t = 4$ (9th order accurate) temporal discretizations have been implemented, and the additional cost for the IMEX schemes has been found to be less than 2% of overall computing cost.

IV. Numerical Results

IV-A. Convection of An Isentropic Vortex

An isentropic vortex convection problem [19, 23, 24] is considered to assess temporal accuracy. As an initial condition, an isentropic vortex is superposed to the mean flow field. Initial mean flow and perturbation values for the isentropic vortex are given by

$$u_\infty = 0.5, v_\infty = 0, p_\infty = \rho_\infty = T_\infty = 1, (\delta u, \delta v) = \frac{\beta}{2\pi} e^{(1-r^2)/2}(-\bar{y}, \bar{x}),\ \delta T = -\frac{(\gamma-1)\beta^2}{8\gamma\pi^2} e^{1-r^2},$$

where β is the vortex intensity set to 4 and $\gamma = 1.4$. Here, $(\bar{x}, \bar{y}) = (x - x_{v0}, y - y_{v0})$, where x_{v0} and y_{v0} are coordinates of the center of initial vortex :$(x_{v0}, y_{v0}) = (0,0)$, and $r^2 = \bar{x}^2 + \bar{y}^2$. The entire flow field is required to be isentropic so, for a perfect gas, $p/\rho^\gamma = 1$.

From the relations, $\rho = \rho_\infty + \delta\rho$, $u = u_\infty + \delta u$, $v = v_\infty + \delta v$, $T = T_\infty + \delta T$, and the isentropic relation, the resulting conservative variables are given by

$$\rho = T^{1/(\gamma-1)} = (T_\infty + \delta T)^{1/(\gamma-1)} = \left[1 - \frac{(\gamma-1)\beta^2}{8\gamma\pi^2} e^{1-r^2}\right]^{1/(\gamma-1)}, \rho u = \rho(u_\infty + \delta u) = \rho\left[1 - \frac{\beta}{2\pi} e^{(1-r^2)/2}\bar{y}\right]$$
$$\rho v = \rho(v_\infty + \delta v) = \rho\left[1 + \frac{\beta}{2\pi} e^{(1-r^2)/2}\bar{x}\right], p = \rho^\gamma, e = \frac{p}{\gamma-1} + \frac{1}{2}\rho(u^2 + v^2).$$

The computational domain is set to $-10 \le x \le 10$ and $-10 \le y \le 10$. Periodic boundary conditions are used in the *x*- and *y*-directions. The vortex convects to the right with the freestream velocity ($u_\infty = 0.5$) and due to the periodic boundary conditions, goes back to the initial location every non-dimensional time interval $\Delta t = 40$.

1) Temporal accuracy study for DGT and FIRK

Since the flowfield is inviscid, the exact solution is just a passive advection of the initial vortex with a mean velocity. The overall error, however, includes both spatial and temporal errors. In order to eliminate the spatial error and to isolate the temporal error, a reference solution for each temporal *p*-order discretization is obtained first using a small time-step of Δt=0.01. Then, the temporal error is computed as the L2 norm of the difference of the density field between the reference solution and a corresponding computed solution.

Table 1 examines the temporal error convergence observed for the 2-stage Gauss FIRK scheme compared with the $p_t = 1$ DGT scheme using Gauss points and the 2-stage Radau FIRK scheme. As seen in the table, the 2-stage Gauss FIRK scheme achieves the design accuracy of $2s = 2p_t + 2$, (i.e. 4th order accurate in this case), while the accuracy of the Gauss DGT scheme is one order lower, and equivalent to the 2-stage Radau IIA FIRK scheme, as expected.

Table 2 examines the same temporal error convergence of the $p_t = 4$ DGT scheme using Gauss points and the 5-stage FIRK Radau IIA scheme using $p_s = 12$ in the spatial dimensions to minimize spatial error. Here as well, the DGT Gauss scheme and the FIRK Radau schemes both achieve equivalent temporal accuracy orders of $2s - 1 = 2p_t + 1$, (i.e. approximately 9th order accuracy is observed on average) as expected. Note that the DGT scheme using right Radau quadrature points was also run and produced results identical to the FIRK Radau IIA scheme in all cases.

Although the DG-in-time scheme using Gauss quadrature points produces results very close to those obtained using right Radau quadrature points, only the latter implementation recovers the exact IRK scheme with its proven accuracy and stability properties. Properties of the DG-in-time scheme using Gauss points are discussed in reference [19] and include A and L stability, similarly to the IRK Radau schemes. However, the use of right-Radau quadrature points obviates the need to compute the projection of the temporal solution from the quadrature (or stage) values to the end of the time interval (using the IRK b_j coefficients), since the last Radau point coincides with the end of the time step interval. Apart from the computational savings afforded by the omission of this last step, this may also lead to better control on the error at the end of the time interval, particularly when only partial convergence of the temporal system is performed, as is most often the case [9].

Table 1 Observed temporal accuracy for convecting vortex problem: p_s=5 and p_t=1 (2-stage)

	Δt	L_1 error	L_1 order	L_2 error	L_2 order	L_∞ error	L_∞ order
2-stage FIRK Gauss	1.000	1.2295E-05		7.3265E-05		1.2441E-03	
	0.500	7.2081E-07	4.09	3.4346E-06	4.41	4.4038E-05	4.82
	0.250	5.3668E-08	3.75	2.5884E-07	3.73	3.8282E-06	3.52
	0.125	3.6222E-09	3.89	1.7224E-08	3.91	2.6195E-07	3.87
	Avg.		3.91		4.02		4.07
DGT with Gauss quadrature points ($p_t = 1$, 2-stage)	1.000	4.6267E-05		2.4704E-04		4.7856E-03	
	0.500	7.5068E-06	2.62	3.9115E-05	2.66	7.6858E-04	2.64
	0.250	1.0173E-06	2.88	4.9508E-06	2.98	9.2563E-05	3.05
	0.125	1.3278E-07	2.94	6.3772E-07	2.96	1.1358E-05	3.03
	Avg.		2.81		2.87		2.91
2-stage FIRK	1.000	3.7273E-05		1.9873E-04		3.7876E-03	

Radau 2A	0.500	5.6780E-06	2.71	3.1314E-05	2.67	6.2362E-04	2.60
	0.250	7.4971E-07	2.92	3.9743E-06	2.98	8.0173E-05	2.96
	0.125	9.5939E-08	2.97	5.0059E-07	2.99	9.9944E-06	3.00
	Avg.		2.87		2.88		2.86

Table 2 Observed temporal accuracy for convecting vortex problem: p_s=12 and p_t=4 (5-stage)

	Δt	L_1 error	L_1 order	L_2 error	L_2 order	L_∞ error	L_∞ order
DGT with Gauss quadrature points (p_t=4, 5 Stage)	1.000	2.4144E-08		1.1703E-07		2.4559E-06	
	0.750	1.9170E-09	8.81	9.8927E-09	8.59	1.8823E-07	8.93
	0.500	4.7182E-11	9.14	2.3433E-10	9.23	3.9927E-09	9.50
	0.250	2.0948E-13	7.82	5.8940E-13	8.64	1.0000E-11	8.64
	Avg.		8.59		8.82		9.02
5-stage FIRK Radau2A	1.000	1.1182E-08		5.2406E-08		1.1918E-06	
	0.750	8.9425E-10	8.78	4.4761E-09	8.55	8.2910E-08	9.27
	0.500	2.1331E-11	9.21	1.0363E-10	9.29	1.7976E-09	9.45
	0.250	1.0938E-13	7.61	4.1460E-13	7.97	7.8826E-12	7.83
	Avg.		8.53		8.60		8.85

For these reasons, in the following sections of this paper, we focus on the use of the Radau IIA IRK scheme, which is identical to the DG-in-time scheme using right Radau quadrature points, although all results have also been reproduced using the DG-in-time Radau scheme.

2) Convergence performance for explicit multi-stage optimized RK

In this section, the convergence performance of the explicit multi-stage optimized RK scheme in pseudo time is investigated. The isentropic vortex convection problem used in the previous section also serves as the time-dependent flow problem. Tests were performed using a fixed spatial discretization ($p_s = 5$) on the domain size of $-10 \leq x \leq 10$ and $-10 \leq y \leq 10$.

The convergence performance is compared in terms of the number of explicit pseudo-time steps required to drive the L2-norm of the space-time residual to less than 10^{-12}. The convergence histories are examined for the solution of a single physical implicit time step at the early time evolution of the vortex test problem. In all cases, it was verified that the convergence behavior was similar at subsequent physical time steps in order to ensure that the observed trends are representative of the general time-dependent solution.

As expected, the solution efficiency depends on the physical time step size, as well as the largest pseudo-time step size that can be used in the iterative scheme without provoking instabilities. The CFL number is an important parameter in this respect. We define a CFL number CFL_t based on the physical time step, as well as a CFL number denoted as CFL_τ based on the pseudo-time step. For time-accurate solutions, the physical-time step is held constant, and the physical-time CFL number is variable, driven principally by the difference cell sizes in the mesh. For pseudo-time stepping, we generally employ local time stepping for convergence acceleration, whereby the pseudo-time CFL number if fixed, resulting in a variable pseudo-time step which depends principally on cell size. A standard definition for the CFL number for a cartesian mesh in two dimensions is given as:

$$CFL_t = (|u| + c)\frac{\Delta t}{min(dx, dy)}$$

where u and c are the local fluid velocity and speed of sound, respectively, and dx, dy represent the cell size in two dimensions. In order to account for the effect of high-order spatial discretizations, where the explicit time step limit becomes more restrictive, we define a modified CFL number as:

$$\mathrm{CFL}_\mathrm{t}^* = (|\mathrm{u}| + \mathrm{c})\frac{\Delta \mathrm{t}}{\min(\mathrm{dx}, \mathrm{dy})}(\mathrm{p}+1)^{1.8}$$

Here p denotes the polynomial order of the spatial discretization. This definition has been derived by empirical numerical evidence [7,20] in an attempt to obtain values of $\mathrm{CFL}_t^* \sim 1$ in the vicinity of the explicit time-step limit for higher order discretizations. For p>0 (i.e. higher than first-order accuracy) a specific value of CFL_t^* corresponds to a smaller time step than that defined by the same value of CFL_t, or conversely, for a given time step size, CFL^*_t is

larger than CFL_t. For example, when $p_s = 5$, CFL_t^* is about 25 times larger than CFL_t. CFL_τ^* for the pseudo-time step is defined in a similar manner. In this work, we employ CFL_t^* and CFL_τ^* as parameters for comparison purposes.

Table 3 shows the maximum stable pseudo-time CFL number (CFL_τ^*) (obtained empirically) and the corresponding number of pseudo-time iterations (ν_τ) to solve the unsteady residual of the DGT/Radau scheme for the isentropic vortex convection problem to a space-time residual tolerance of 10^{-12}. The notation RK*sp* indicates an *s*-stage scheme with *p*-order of temporal accuracy. RK_{11}, RK_{44} and RK_{41} correspond to single-stage RK, 4-stage standard RK (4^{th} order accurate in τ) and 4-stage optimized RK (1^{st} order accurate in τ), respectively. The number of stages (*s*) is used to normalize $CFL*_\tau$ and ν_τ for comparison purposes between single- and multi-stage schemes in the table. In order to assess the effect of the physical-time step size Δt on convergence performance of the optimized RK schemes, the physical-time steps of Δt=0.26 and 2.6, which correspond to $CFL_t^* = 25$ and 250 respectively (based on a uniform mesh cell size and freestream conditions), are used for comparison. Results are given for DGT physical temporal discretizations ranging from p_t=1 to p_t=5, corresponding to the IRK Radau scheme using from 2 to 6 stages. The number of cycles to convergence multipled by the number of stages (i.e. residual evaluations) is directly related to the maximum stable CFL_τ/s value. The small physical time step problems all converge faster than the larger time-step problems, as expected. Interestingly, relatively similar convergence rates are observed overall for increasing orders of temporal accuracy or p_t values. The relative speedups obtained by the RK_{44} and RK_{41} schemes are obtained from the results in Table 3 and displayed in Table 4. Compared to RK_{11}, RK_{44} shows 0.61x~1.6x and 2.5x~4.3x speed-up factors for Δt=0.26 and 2.6, respectively. The speed-up factors increase for higher-order p_t values of the temporal discretization. RK_{41} shows a 1.8x speed-up factor compared to RK_{44} as predicted and verified in reference [17]. The overall speedup going from the simple RK11 scheme to the optimized RK41 scheme for the larger physical time step problem is in the range of 7x to 8x for all temporal accuracy orders above p_t=1 or IRK Radau schemes with 3 to 6 stages.

Table 3 Comparison of pseudo-time CFL number and corresponding number of residual evaluations (i.e. pseudo-time iterations multiplied by number of stages) to solve unsteady residual of DGT/Radau implicit time stepping to a residual level of 1.e-12 for physical time step of 0.26 and 2.6

	Δt=0.26						Δt=2.6					
	RK_{11}		RK_{44}		RK_{41}		RK_{11}		RK_{44}		RK_{41}	
p_t	CFL_τ^*/s	$\nu_\tau * s$	CFL_τ^*/s	$\nu_\tau * s$	CFL_τ^*/s	$\nu_\tau * s$	CFL_τ^*/s	$\nu_\tau * s$	CFL_τ^*/s	$\nu_\tau * s$	CFL_τ^*/s	$\nu_\tau * s$
1.0	1.2	80	0.75	132	1.38	72	0.29	3296	0.70	1332	1.28	736
2.0	1	128	0.88	132	1.58	76	0.18	6501	0.78	1452	1.45	776
3.0	1.1	133	1.18	120	2.15	64	0.29	4762	1.08	1248	1.98	680
4.0	1.1	156	1.48	108	2.45	64	0.29	5374	1.23	1240	2.25	676
5.0	1.2	177	1.7	108	2.68	64	0.36	4830	1.50	1136	2.75	616

Table 4 Comparison of speed-up factors for physical time step of 0.26 and 2.6

	Δt=0.26			Δt=2.6		
p_t	RK_{11}/RK_{44}	RK_{44}/RK_{41}	RK_{11}/RK_{41}	RK_{11}/RK_{44}	RK_{44}/RK_{41}	RK_{11}/RK_{41}
1.0	0.61	1.83	1.11	2.47	1.81	4.48
2.0	0.97	1.74	1.68	4.48	1.87	8.38
3.0	1.11	1.88	2.08	3.82	1.84	7.00
4.0	1.44	1.69	2.44	4.33	1.83	7.95
5.0	1.64	1.69	2.77	4.25	1.84	7.84

3) Convergence performance for IMEX multi-stage optimized RK

In this section we assess the convergence performance of the IMEX multi-stage optimized RK scheme in pseudo time for the isentropic vortex convection problem. In the previous section, the physical time steps were chosen to be large enough that the IMEX scheme was not required for good convergence behavior. In this section, we focus on a range of physical time steps that extends into the region of small time steps where the IMEX scheme should prove to be beneficial.

Figure 1 shows the convergence histories of IMEX-RK41 for the DGT/IRK Radau temporal discretization with $\Delta t=2x10^{-4}$, which corresponds to a physical CFL number of $\mathrm{CFL}_t^* = 0.02$, based on a uniform mesh cell size and freestream conditions. Temporal orders of accuracy up to $p_t = 4$ were tested, corresponding to 5 stages in the physical time IRK Radau scheme. Green and blue lines indicate the best convergence rate cases for the explicit RK41 and IMEX-RK41 schemes, respectively. For this small physical time step, the explicit RK41 scheme requires a very small CFL_τ^*. The largest stable value of $CFL_\tau^*= 0.1$ is seen to converge slowly in the $p_t=1$ case, in Figure 1(a). Conversely, the IMEX-RK41 scheme converges much faster using this same CFL^*_τ value. However, using a lower value of $CFL^*_\tau= 0.06$ produces the fastest convergence for the explicit RK41 scheme in pseudo time. On the other hand, for very small physical time steps, the physical temporal term dominates and the IMEX scheme becomes fully implicit in pseudo time, enabling the use of very large CFL^*_τ values, which provides a large speedup over the best explicit RK41 performance. The relative speedup of the best IMEX RK41 performance versus the explicit RK41 scheme increases as the temporal order of accuracy of the physical time discretization is raised as seen in Figures 1 (a) through (d). Notably at the higher temporal orders, the convergence of the IMEX scheme using CFL^*_τ values that are close to the stability limit of the explicit scheme is still faster than the best convergence obtained with the explicit scheme.

The best performance in terms of the number of sub-iterations for the IMEX versus explicit scheme is compared in Table 5. Compared to the explicit RK41 scheme, IMEX-RK41 shows 2.0x~8.7x speed-up for this small physical time step problem. The speed-up factors increase for higher-order cases (i.e. larger number of stages in physical time). Similar tests were made for larger physical time steps of $\Delta t=0.02$ ($\mathrm{CFL}_t^* = 2$) and 0.26 ($\mathrm{CFL}_t^* = 25$). The results are presented in Table 6 and Table 7. Similarly to Table 5, the speed-up factors increase as p_t increases, although these are more modest at 1.4x~4.3x and 1.0x~1.6x for $\Delta t=0.02$ and 0.26, respectively. As the physical time step size increase, the IMEX-RK41 scheme becomes equivalent to the explicit RK41 pseudo-time stepping scheme, which is expected since the temporal residual term goes to zero when Δt is large. Although convergence of the IMEX scheme in pseudo time can be enhanced by choosing the optimal CFL^*_τ value depending on the size of the physical time step, an important aspect of the IMEX scheme is that it remains stable for all physical time steps using a fixed CFL^*_τ value of O(1), as can be seen from the results of these tables, thus simplifying the implementation of pseudo-time stepping schemes for problems with variable physical time steps and mesh cells of widely varying resolution.

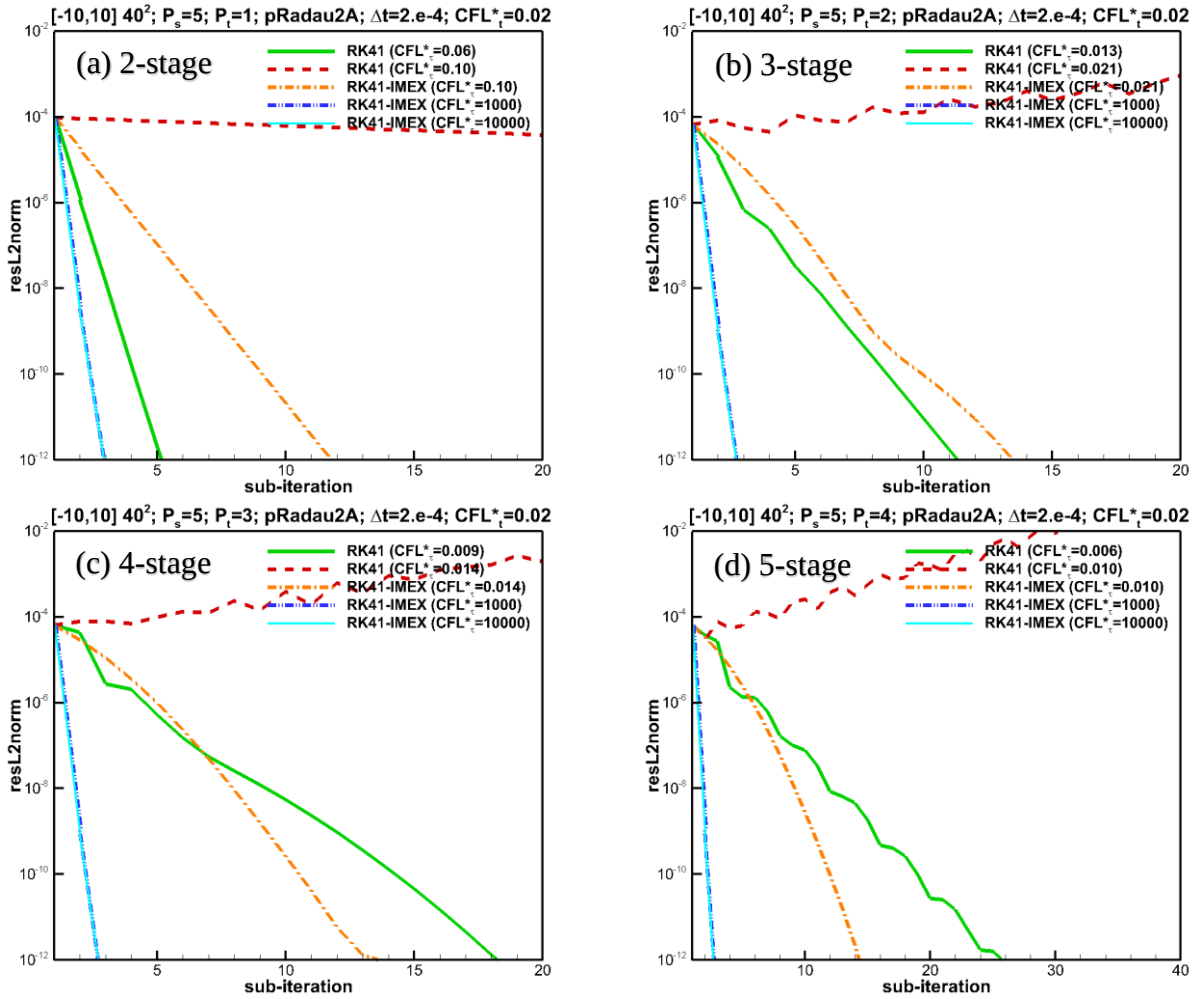


Figure 1: *Convergence histories of IMEX-RK41 for preconditioned Radau 2A (pFIRK-R) with* Δt*=0.0002: (a)* p_t*=1, (b)* p_t*=2, (c)* p_t*=3, (d)* p_t*=4*

Table 5: *Best performance in terms of the number of sub-iterations for pFIRK-R with* Δt*=0.0002*

		explicit RK41		IMEX-RK41		explicit/IMEX
# of stages	p_t	CFL^*_τ	# of sub-iteration	CFL^*_τ	# of sub-iteration	speed-up factor
2	1	0.060	6	1000	3	2.0
3	2	0.013	12	1000	3	4.0
4	3	0.009	19	1000	3	6.3
5	4	0.006	26	1000	3	8.7

Table 6: *Best performance in terms of the number of sub-iterations for pFIRK-R with* Δt*=0.02*

		explicit RK41		IMEX-RK41		explicit/IMEX
# of stages	p_t	CFL^*_τ	# of sub-iteration	CFL^*_τ	# of sub-iteration	speed-up factor
2	1	2.0	20	7.0	14	1.4
3	2	1.1	13	30	6	2.2
4	3	0.8	19	100	6	3.2
5	4	0.6	26	100	6	4.3

Table 7: *Best performance in terms of the number of sub-iterations for pFIRK-R with* Δt*=0.26*

		explicit RK41		IMEX-RK41		explicit/IMEX
# of stages	p_t	CFL_τ^*	# of sub-iteration	CFL_τ^*	# of sub-iteration	speed-up factor
2	1	2.5	218	2.6	225	1.0
3	2	2.4	82	3.0	65	1.3
4	3	2.3	73	3.2	47	1.6
5	4	2.3	64	3.3	40	1.6

IV -B. Flows Around 2D/3D NACA0012 Airfoils

Although the previous test case was instrumental in demonstrating the advantages of the optimized RK and IMEX schemes in pseudo time, the mesh used for the 2D vortex problem is essentially uniform in space. One of the principal motivations for using implicit time-stepping schemes for scale-resolving simulations is the need to accommodate meshes of widely varying resolution, where explicit schemes become constrained by the time-step restrictions determined by the smallest cells in the mesh. The use of implicit schemes with a fixed global physical time step on meshes of widely varying resolution can result in regions of the domain operating at very large physical time-step CFL numbers, well beyond the stability limit of explicit schemes in small cell regions, while large cell regions may operate at very low CFL numbers that may be near or below the explicit time-step limit. In this sense, the IMEX scheme can be expected to be instrumental in providing consistent convergence properties throughout the domain, while the use of local pseudo-time stepping should prove to be effective in accelerating the solution of the implicit system at each physical time step.

In this section we examine the performance of these schemes for 2D and 3D flow over a NACA 0012 airfoil using a body fitted mesh with large variation in cell sizes. The Mach number is 0.2, the Reynolds number is 60,000 and the angle of attack is 8°, which is the same flow condition used in [17]. Laminar separation occurs on the upper surface of the airfoil, forming a laminar separation bubble. The DGT and IRK temporal discretizations have been implemented for curved meshes in 2D and 3D. The DG4est code [7,20] has been extended to 3D unstructured curved element meshes, incorporating dynamic *hp* refinement with the p4est library. All high-order temporal discretizations introduced in the previous section have also been implemented in this 3D code. In the current test cases, DGT/Radau with $p_t = 1{\sim}4$ is used without any *hp* refinement for simplicity. $p_s = 2$ or third order spatial accuracy is used for these simulations.

1) 2D NACA0012 airfoil

The grid employed for the 2D NACA0012 airfoil and the u-velocity iso-surfaces for a representative 3D simulation on are shown in Figure 2. The grid contains a total of 9,633 curved quadratic (Q2) quadrilateral cells with extensive refinement in the leading edge and trailing edge regions, and extends out to 100 chord lengths in the farfield. In this test, the 2D flow field over the NACA0012 airfoil starts developing at $t = 0$ from an initial freestream field (here t indicates physical time). Then, we choose a specific physical time location $t = t_s$ and restart the flow solver with a physical time step size Δt. The convergence histories at $t = t_s$ with Δt are compared to assess the convergence performance of the IMEX scheme. Two physical time locations t_s=2.e-3 and 10 are considered to see if there is any noticeable change in convergence performance as the flow develops a laminar separation bubble.

Figure 3 shows convergence histories at t =2.e-3 with Δt=2.e-3 ($CFL_t^* = 100$), i.e. at the second time step after startup. Convergence of the implicit system for the DGT/Radau physical temporal discretization for p_t=1,2,3,4 is examined. Because of the large variation in cell size, $\Delta\tau_{max}/\Delta\tau_{min}$ for a given CFL_τ^* is of the order of 100,000 on this mesh. The given physical time step size of Δt=2.e-3 corresponds to a physical CFL number of approximately CFL^*_t=100, for the smallest cell in the mesh, based on freestream conditions. This results in physical time CFL values several orders of magnitude below 1 for the larger cells in the mesh, where the temporal term in the space-time residual becomes dominant. For the purely explicit pseudo-time stepping scheme, this requires a significant reduction in the pseudo-time CFL number in order to maintain stability in the large cells. For example, for the p_t=1 case, the pseudo-time CFL number must be reduced to 0.006 to maintain stability, resulting in very slow convergence. When the IMEX scheme is run with this same CFL value, equivalently slow convergence is observed. However, the IMEX scheme remains stable for pseudo-time CFL values greater than 1, for all cases, due to the implicit treatment of the temporal part of the space-time residual. This results in fast and consistent convergence, achieving full convergence (to 1.e-12) in several hundred pseudo-time steps. Additionally, the IMEX-RK41 scheme shows slightly faster convergence at higher temporal orders of accuracy, with the p_t=4 case delivering approximately 1.6x faster convergence than the p_t=1 case.

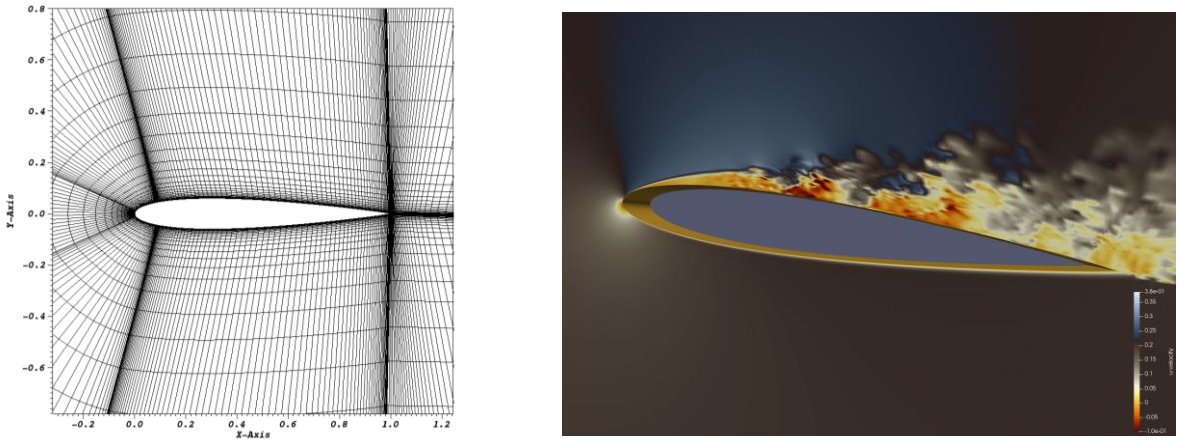


Figure 2 *Grid and u-velocity contour for 2D/3D NACA0012 airfoil simulations: (left) grid for 2D NACA0012 airfoil, (right) u-velocity isosurfaces for 3D NACA0012 airfoil*

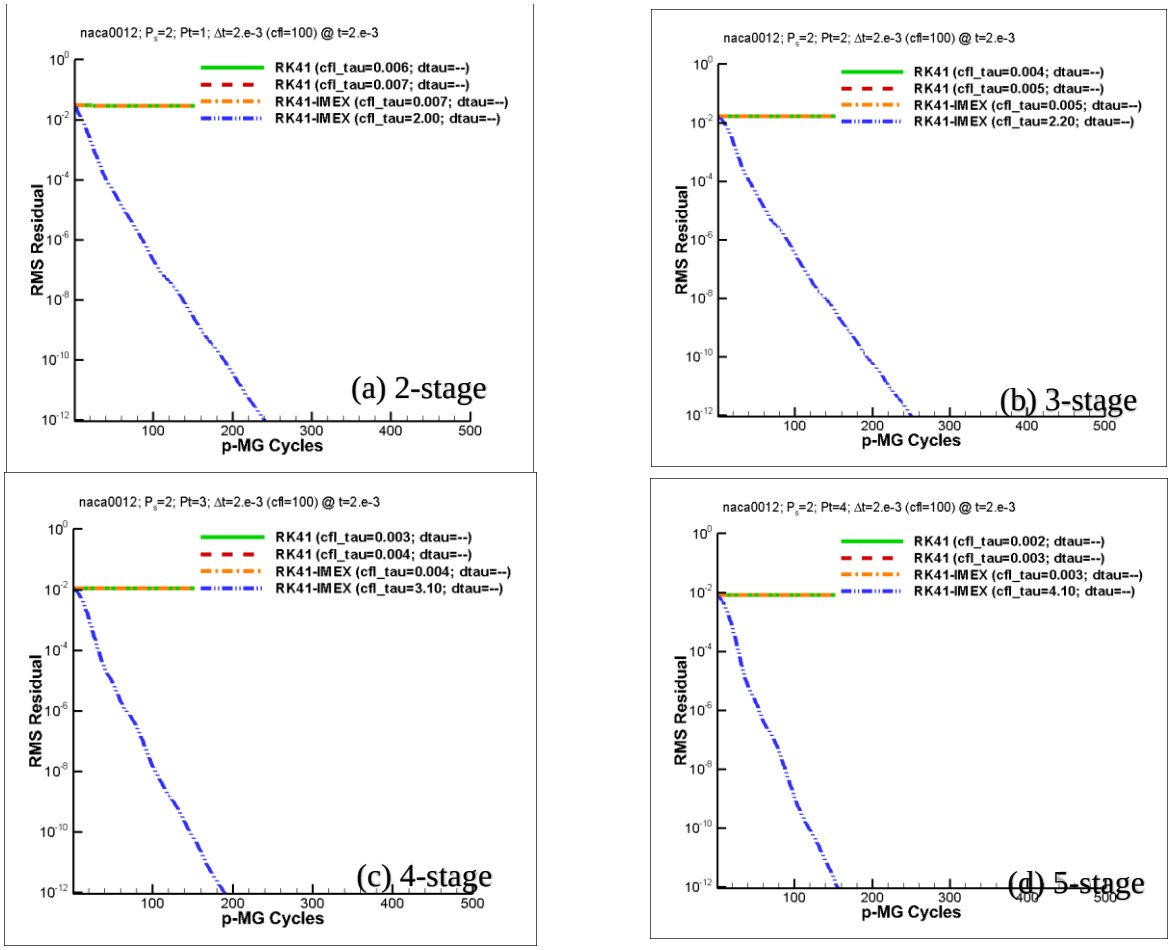
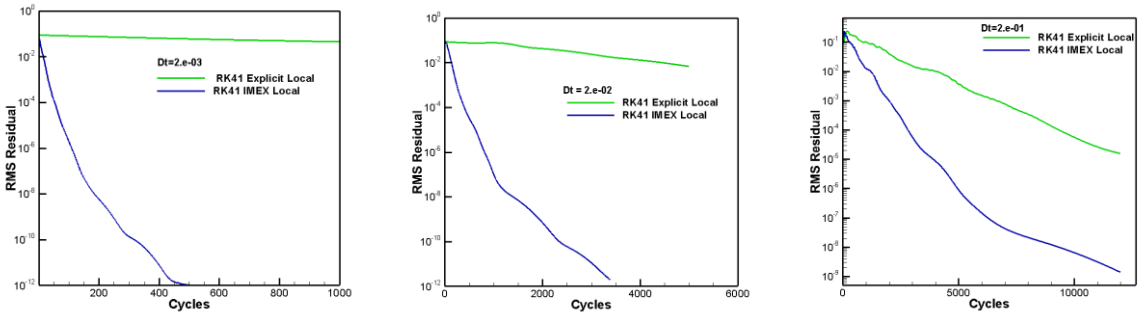


Figure 3 *Convergence histories for 2D NACA0012 airfoil at the second time step after startup for Δt=2.e-3: (a)* $\boldsymbol{p_t}$*=1, (b)* $\boldsymbol{p_t}$*=2, (c)* $\boldsymbol{p_t}$*=3, (d)* $\boldsymbol{p_t}$*=4*



(a) (b) (c)

Figure 4 *Convergence histories for 2D NACA0012 airfoil at t=10.0: (left) Δt=2.e-3, (middle) Δt=2.e-2, (right) Δt=2.e-1 corresponding to CFL^*_t values for the smallest cell based on freestream values of approximately 100, 1000 and 10,000 respectively.*

Figure 4 shows convergence histories at t=10.0 when a laminar separation bubble is in the process of formation. Here we examine the relative performance of the explicit and IMEX RK41 schemes as a function of the physical time step size for a p_t=1 (2-stage) temporal discretization with p_s=2 (third-order) spatial discretization. All cases represent the best performance achieved by each scheme using the highest stable pseudo-time CFL number as determined empirically. For the smallest time step of Δt=2.e-3 (corresponding to CFL^*_t=100 at the smallest mesh cell), the convergence results are similar to those presented previously, although convergence of the IMEX scheme is obtained in 450 cycles, indicating that the implicit system is slightly stiffer at this later simulation time than in the previous case. Similarly to the previous case, the explicit RK41 scheme in pseudo time required a CFL^*_τ value of 0.006 for stability. For the larger physical time step cases of Δt=2.e-2 and Δt=2.e-1, corresponding to smallest mesh-cell physical time CFL values of 1,000 and 10,000 respectively, the maximum stable pseudo-time CFL value for the explicit scheme increased to 0.06 and 0.6, respectively. On the other hand, the IMEX scheme maintains the same optimal CFL^*_τ= 2.0 for all physical time step sizes. However, convergence of the IMEX scheme becomes slower as the physical time step size increases due to the increased stiffness of the implicit system to be solved. Overall, as the physical time step is increased, the performances of the IMEX and explicit RK41 schemes grow closer to each other, as expected.

2) 3D NACA0012 airfoil

A set of fully three-dimensional simulations for a NACA0012 wing section of span 20% chord have been run in order to examine the performance of the explicit and IMEX RK41 schemes in three dimensions. The mesh contains a total of 73,242 quadratically curved (Q2) hexahedral elements. The variation of cell size for this mesh is more moderate than that of the 2D mesh described previously, resulting in a $\Delta\tau_{max}/\Delta\tau_{min}$ of approximately 1,000 for a fixed CFL^*_τ value. Convergence in pseudo-time is examined for two physical time step sizes applied at the physical simulation time of t_s=10. The first time-step size is Δt=2.e-05, corresponding to CFL^*_t=1.36 for the smallest mesh cell, while the second time step size is Δt=2.e-03, (CFL^*_t=136), which is similar to the first case examined in 2D, previously. Figure 5 shows convergence histories for the 3D NACA0012 airfoil case with Δt=2.e-5 ($CFL^*_t = 1.36$). This time step size results in a physical CFL value that is close to or less than 1 at all cells in the mesh. As expected, the explicit RK41 scheme requires a small CFL^*_τ value to maintain stability and converges very slowly. On the other hand, the IMEX scheme becomes unconditionally stable and converges in just several steps using a large CFL^*_τ value of 100 or higher. This suggests that for small physical time steps, the IMEX scheme remains competitive with a purely explicit scheme in physical time, as is discussed further in the next section.

Figure 6 shows the convergence histories obtained for the Δt=2.e-3 ($CFL^*_t = 136$) case. Here the IMEX scheme remains significantly faster than the explicit RK41 scheme, although the difference between these two schemes is less pronounced than in the smaller physical time step case, as expected. Comparing Figure 6 to the results using the same time step size in the 2D case from Figure 3 shows that the performance of the IMEX and explicit RK41 schemes are closer together in the 3D case than in the 2D case for the same physical time step size. This is due to the

more moderate variation of the cell sizes in the 3D grid compared to the 2D grid, which enables larger stable CFL^*_τ values in the latter case. In Figure 6, the IMEX scheme is seen to converge consistently in 100 to 150 steps for all temporal discretizations of p_t=1,2,3,4. By comparison, our purely explicit 4-stage RK scheme in physical time operates at a maximum CFL value of 0.5 [7,20], which would require 270 steps to cover the equivalent time increment of Δt=2.e-03. This suggests that the IMEX scheme can be competitive with the explicit scheme in physical time on this mesh even for these relatively small physical time step sizes.

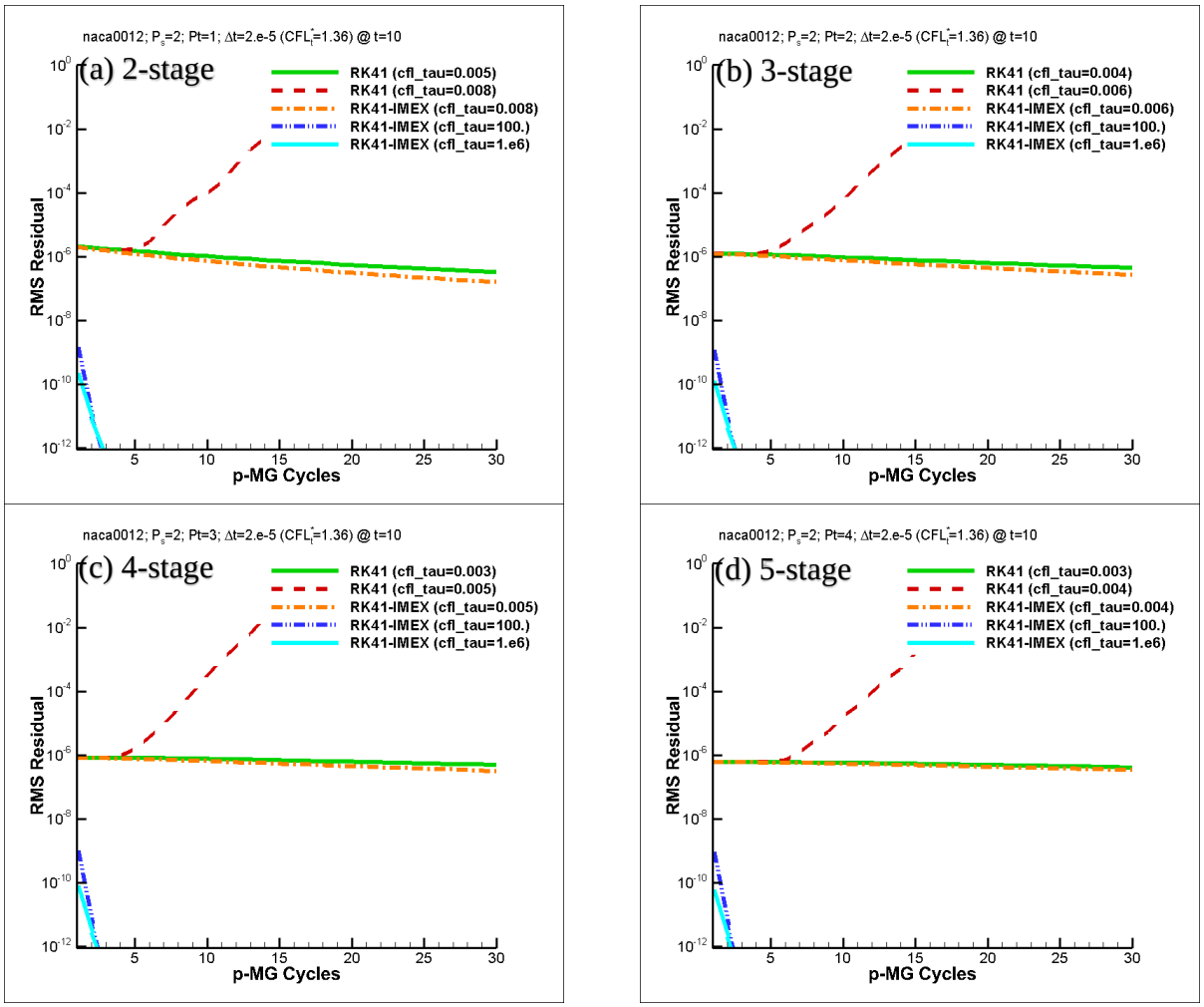


Figure 5: *Convergence histories for 3D NACA0012 airfoil with Δt=2.e-5: (a) $\boldsymbol{p_t}$=1, (b) $\boldsymbol{p_t}$=2, (c) $\boldsymbol{p_t}$=3, (d) $\boldsymbol{p_t}$=4*

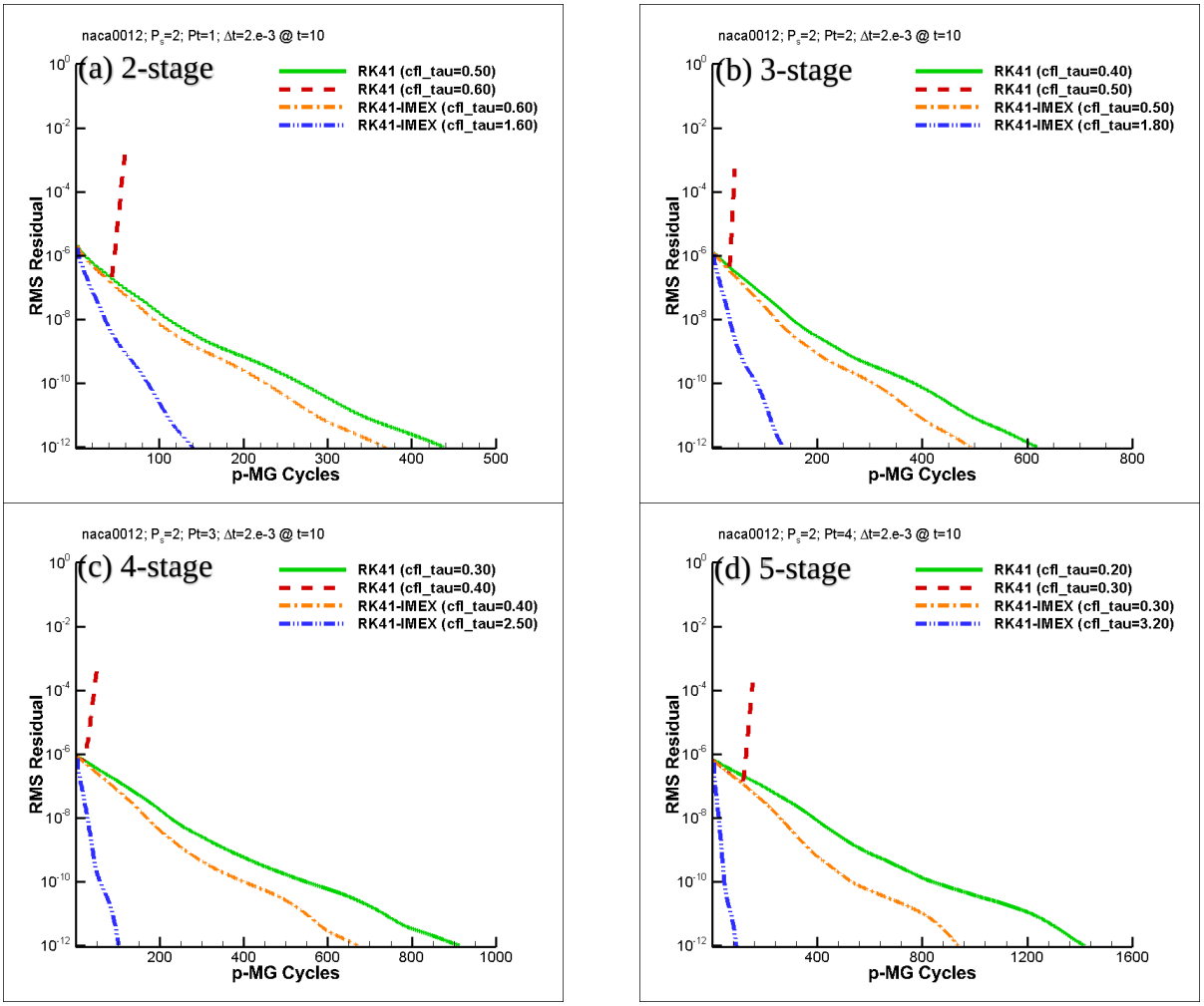


Figure 6: *Convergence histories for 3D NACA0012 airfoil with Δt=2.e-3: (a)* $\boldsymbol{p_t}$*=1, (b)* $\boldsymbol{p_t}$*=2, (c)* $\boldsymbol{p_t}$*=3, (d)* $\boldsymbol{p_t = 4}$

4) Effect of Local versus Global Pseudo-Time Stepping

In all the results discussed above, local time stepping has been used in pseudo time for both the IMEX and explicit RK schemes, whereby a fixed CFL number in pseudo time is used, resulting in a locally variable pseudo time step throughout the mesh. However, as discussed previously, for small to moderate physical time steps, in the case of the explicit RK scheme in pseudo time, the maximum stable CFL^*_τ value is often determined by the larger cells in the mesh, where the temporal term in the space-time residual becomes dominant due to the relatively small (i.e. < 1) physical CFL number, resulting in slow convergence. Perhaps counter-intuitively, the convergence of the explicit RK scheme can be significantly accelerated by abandoning local time stepping and using a global minimum pseudo time step. This is illustrated in Figure 7, which reproduces the 2D results of Figure 4, with the addition of the convergence histories using a global minimum pseudo time step for both the explicit and IMEX RK41 schemes. As shown in Figure 7, for the smallest physical time step of Δt=2.e-03, the explicit RK41 scheme using local time stepping that was shown to be very slow, becomes competitive with the IMEX scheme when switching to global time stepping in pseudo time. Additionally, the use of global pseudo-time stepping with the IMEX scheme produces almost identical results to the explicit scheme with global pseudo-time stepping.

To examine the behavior of this approach, we consider the pseudo-time stepping equations for a simple BDF1 temporal discretization using a single stage RK1 (forward Euler) time-stepping scheme in pseudo time, given previously by equation (19) for the explicit scheme, and by equation (20) for the implicit scheme. In the explicit scheme, for small physical time steps, the second term on the right-hand side in equation (19) can become dominant and cause instabilities. While the value of the physical time step Δt is fixed, for local pseudo time-stepping, the

value of $\Delta\tau$ increases with cell size (at fixed CFL^*_τ), adding to the growth and instability caused by the temporal term in equation (19). For this reason, the maximum stable CFL^*_τ for local pseudo-time stepping is often set by the largest cells in the mesh and becomes very restrictive. On the other hand, using a (constant) global minimum pseudo-time step results in a constant ratio $\Delta\tau/\Delta t = \alpha$ throughout the entire mesh. In this case, equation (19) can be written as:

$$Q^{k+1} = (1 - \alpha)\, Q^k + \alpha Q^n - \alpha \Delta t\, M^{-1}\, R\left(Q^k\right)$$

If the physical time CFL number is close to 1 at the smallest cell in the mesh, the global minimum pseudo time step will be close to the physical time step and $\alpha \sim 1$. In this case the above equation reduces to:

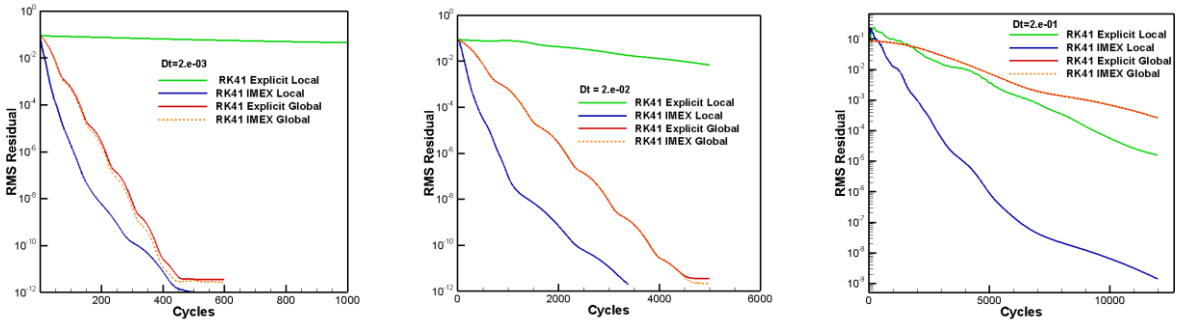
$$Q^{k+1} = Q^n - \Delta t M^{-1} R(Q^k)$$

This shows that in such a case, global minimum pseudo-time stepping reduces to explicit time stepping in physical time.

For the implicit pseudo-time scheme described by equation (20) (representative of the IMEX RK scheme), for small physical time steps, the factor on the left-hand side can be approximated as $(1 + \Delta\tau\backslash\Delta t) \sim \Delta\tau/\Delta t$. In this case equation (19) can be rewritten as:

$$Q^{k+1} = \frac{\Delta t}{\Delta \tau} Q^k + Q^n - \Delta t M^{-1} R(Q^k)$$

When the ratio $\Delta\tau/\Delta t$ is large, the first term on the right-hand side, which contains the inverse ratio may be neglected, showing how the implicit scheme with local pseudo-time stepping takes on the same form as the explicit scheme with global minimum pseudo-time stepping, and reduces to explicit time-stepping in physical time. This equivalence explains the dramatic improvement in convergence behavior of the explicit scheme using global minimum pseudo-time stepping and the similar convergence characteristics of this approach and the IMEX scheme in Figure 7(a). This may explain the use of global pseudo-time stepping in the literature, for example in one of the cases reported in reference [17].



(a) (b) (c)

Figure 7: *Comparison of local and global minimum pseudo-time stepping schemes for 2D NACA0012 airfoil at t=10.0: (a) Δt=2.e-3, (b) Δt=2.e-2, (c) Δt=2.e-1 corresponding to CFL^*_t values for the smallest cell based on freestream values of approximately 100, 1000 and 10,000 respectively.*

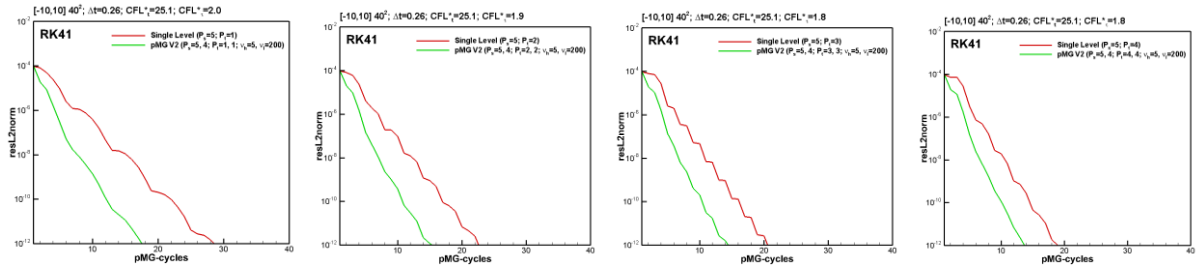
However, as the physical time step increases and the temporal term in the space-time residual becomes less important, the global minimum pseudo-time stepping approach remains constrained to operate at fixed pseudo-time step value $\Delta\tau = \alpha\, \Delta t$ (determined by the smallest cell) which is much lower than the pseudo-time explicit stability limit in the larger cells of the domain. This explains the relative slowdown of the explicit global pseudo-time stepping approach with increasing physical time-step size in Figure 7. For a physical time-step size of Δt=2.e-02 (corresponding to a minimum cell CFL^*_t=1000), the explicit global-minimum pseudo-time stepping scheme becomes significantly slower than the IMEX scheme, although it is still faster than the explicit local-time stepping scheme. However, for the largest physical time step size of Δt=2.e-01 (corresponding to a minimum cell

CFL^*_t=10,000) the explicit global-minimum pseudo-time stepping scheme becomes slower than all other schemes, as expected. The relative performance of these schemes depends strongly on the physical time step and the variation of cell sizes within the computational mesh. However, in all cases the IMEX scheme with local pseudo-time stepping produces the fastest convergence rates. The ability to operate near the local stability limit in pseudo-time is important also for multigrid methods, in order to obtain the design error smoothing properties of the scheme on the fine mesh level, which are instrumental for the efficiency of the multigrid algorithm.

V. P-Multigrid Acceleration

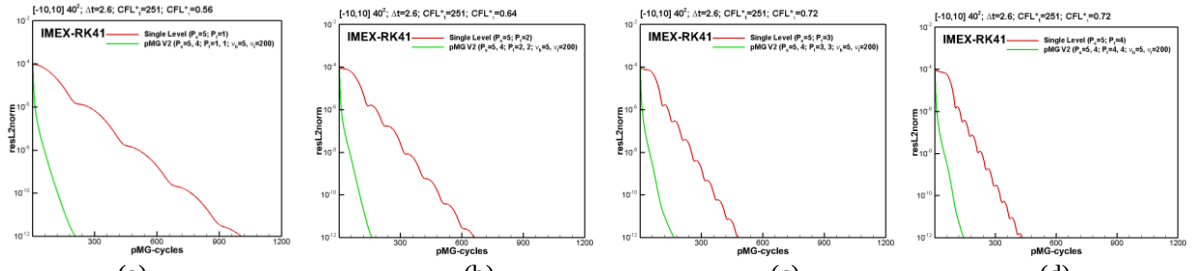
As demonstrated in Figure 7 in the previous section, the convergence of the pseudo-time stepping approach slows down significantly as the physical time step is increased, due to the increased stiffness of the implicit system of equations generated by large physical time-step problems. One approach to accelerate the convergence of the implicit solver is to use the pseudo-time stepping approach as a smoother on the different levels of a multigrid scheme. For high-order tensor-product space-time discretizations, this approach retains the efficiency of pseudo-time stepping where the use of Jacobian matrices is avoided and only residual evaluations are required. Based on the results of the previous sections, we make use of the IMEX RK41 scheme with local pseudo-time stepping exclusively as a smoother, since this scheme was shown to provide the best performance overall as a single grid solver. For simplicity we return to the 2D vortex convection problem discussed in Section IV.A, and examine the convergence of a two-level p-multigrid solver, for DGT/Radau temporal discretizations of orders p_t=1,2,3,4 using p_s=5 (6^{th} order accuracy in space). Here the coarse multigrid level consists of a p_s=4 (5^{th} order accuracy in space) discretization with the same temporal discretization as the underlying fine mesh. We perform 5 pseudo time-steps on the fine level, followed by 200 pseudo time steps on the coarser level, in an attempt to adequately solve the coarse grid problem at each p-multigrid cycle. These 2-level multigrid schemes are not computationally efficient due to the expense of solving the coarse grid problem. However, they provide an estimate of the optimum convergence in terms of p-multigrid cycles achievable for the given fine grid smoother and inter-grid transfer operations.

Figure 8 compares the convergence rates of the single grid and 2-level multigrid solver for different temporal orders of accuracy using a physical time step size of Δt=0.26, which corresponds to a physical time CFL number of CFL^*_t=25.1, while Figure 9 provides the same comparisons for a physical time step size of Δt=2.6, corresponding to a physical time CFL number of CFL^*_t=251.



(a) (b) (c) (d)

Figure 8: *Comparison of single versus 2-level p-multigrid solver for 2D vortex convection test case for a physical time step of Δt=0.26 for (a) p_t=1, (b) p_t=2, (c) p_t=3 and (d) p_t=4 temporal discretizations.*



(a) (b) (c) (d)

Figure 9: *Comparison of single versus 2-level p-multigrid solver for 2D vortex convection test case for a physical time step of Δt=2.6 for (a) p_t=1, (b) p_t=2, (c) p_t=3 and (d) p_t=4 temporal discretizations.*

The multigrid solver provides a significant speedup in terms of cycles for the larger time-step case, with more moderate acceleration being observed for the smaller time-step case. Table 8 depicts the speed up obtained for the larger time-step case, showing acceleration factors ranging from 4.8 for the p_t=1 temporal discretization to 3.0 for the p_t=4 discretization. Larger speedups can be expected for larger physical time-step cases, as the stiffness of the implicit system increases with increasing physical time-step sizes. These results demonstrate that the IMEX RK41 scheme remains effective as a fine level multigrid smoother over a range of physical time step sizes and for low to high-order accurate DGT/Radau temporal discretizations. However, in order to construct a cost-effective multigrid solver, a suitable coarse grid solver that operates on a recursive sequence of coarser meshes must be developed. Further improvements may be sought by developing an RK pseudo-time stepping scheme optimized for high frequency error damping (rather than explicit time step size as in [17]) and invoking temporal coarsening for high p_t discretizations.

Table 8: *Number of p-multigrid cycles for single level and 2-level p-multigrid solver for physical time step of Δt=2.6 and associated speedup factors at different temporal orders of accuracy.*

		IMEX-RK41 # of cycles		
pt	cfl_tau	single	MG V2	speed-up factor
1	0.56	1007	208	4.8
2	0.64	660	162	4.1
3	0.72	476	166	2.9
4	0.72	433	144	3.0

A. Conclusion and Future Work

In this work we have considered both DG in time and IRK temporal discretizations and clarified the relationships between these approaches as well as their stability and accuracy properties, based on results from references [10,19]. We have developed an efficient IMEX pseudo-time stepping scheme for solving the implicit systems that arise from space-time DG discretizations for varying orders of accuracy in space and time. The IMEX scheme is based on an optimized RK41 local pseudo-time stepping scheme that was shown to deliver a speedup of 1.8 over the traditional 4-stage fourth-order accurate explicit RK scheme, with up to 7.8x speedup demonstrated over a simple single stage explicit approach in pseudo-time. The IMEX RK41 scheme was shown to deliver efficient and robust convergence rates for space-time DG discretizations over a wide range of temporal orders of accuracy, physical time step sizes and for meshes with large variations in cell sizes in both two dimensions and three dimensions. The proposed IMEX scheme is instrumental for retaining fast convergence rates for moderate size physical time step problems and was also shown to mimic explicit time-stepping in physical time in the presence of small physical time steps. This work has also clarified the effect of local versus global pseudo-time stepping on meshes with wide ranges of cell sizes. Although the poor convergence properties of the corresponding explicit RK pseudo-time stepping scheme in the presence of small physical time steps can be rectified through the use of global minimum pseudo-time stepping, this approach was also shown to degrade in the presence of larger physical time steps on meshes with large variations of cell sizes. The proposed IMEX scheme was also shown to be suitable for use as a smoother in a p-multigrid solver, with the potential to accelerate the convergence of DG-in-time discretizations with large physical time steps. The IMEX scheme is well suited for use with scale-resolving methods, where it can remain competitive with explicit time-stepping approaches in physical time for small time steps, while at the same time mitigate the effect of overly restrictive explicit stability limits due to wide variations in mesh cell size. Future work will focus on the development of an optimal IMEX RK scheme for p-multigrid efficiency and demonstration in more complex three-dimensional scale-resolving simulations.

Acknowledgments

This work was supported by NASA Grant 80NSSC18M0154 under the T^3 project.

References

[1] J. J. W. van der Vegt and H. van der Ven, “Space–Time Discontinuous Galerkin Finite Element Method with Dynamic Grid Motion for Inviscid Compressible Flows: I. General Formulation,” *J. Comput. Phys.*, vol. 182, no. 2, pp. 546–585, Nov. 2002, doi: 10.1006/jcph.2002.7185.

[2] C. M. Klaij, J. J. W. van der Vegt, and H. van der Ven, “Space–time discontinuous Galerkin method for the compressible Navier–Stokes equations,” *J. Comput. Phys.*, vol. 217, no. 2, pp. 589–611, Sep. 2006, doi: 10.1016/j.jcp.2006.01.018.

[3] L. T. Diosady and S. M. Murman, “Tensor-product preconditioners for higher-order space–time discontinuous Galerkin methods,” *J. Comput. Phys.*, vol. 330, pp. 296–318, Feb. 2017, doi: 10.1016/j.jcp.2016.11.022.

[4] W. Pazner and P.-O. Persson, “Approximate tensor-product preconditioners for very high order discontinuous Galerkin methods,” *J. Comput. Phys.*, vol. 354, pp. 344–369, Feb. 2018, doi: 10.1016/j.jcp.2017.10.030.

[5] L. T. Diosady and S. M. Murman, “Scalable tensor-product preconditioners for high-order finite-element methods: Scalar equations,” *J. Comput. Phys.*, vol. 394, pp. 759–776, Oct. 2019, doi: 10.1016/j.jcp.2019.04.047.

[6] S.-H. Yoon and D. Mavriplis, “High-Order Space-Time Discontinuous Galerkin Discretization Using Tensor-Product Formulations,” in *AIAA Scitech 2021 Forum*, 0 vols., American Institute of Aeronautics and Astronautics, 2021. doi: 10.2514/6.2021-0156.

[7] A. Kirby, “Enabling High-Order Methods for Extreme-Scale Simulations,” University of Wyoming, 2018.

[8] D. J. Mavriplis and C. R. Nastase, “On the geometric conservation law for high-order discontinuous Galerkin discretizations on dynamically deforming meshes,” *J. Comput. Phys.*, vol. 230, no. 11, pp. 4285–4300, May 2011, doi: 10.1016/j.jcp.2011.01.022.

[9] A. Jameson, “Evaluation of Fully Implicit Runge Kutta Schemes for Unsteady Flow Calculations,” *J. Sci. Comput.*, vol. 73, no. 2, pp. 819–852, Dec. 2017, doi: 10.1007/s10915-017-0476-x.

[10] H. T. Huynh, “Collocation and Galerkin Time-Stepping Methods,” in *19th AIAA Computational Fluid Dynamics*, 0 vols., American Institute of Aeronautics and Astronautics, 2009. doi: 10.2514/6.2009-4323.

[11] J. D. Lambert, *Numerical Methods for Ordinary Differential Systems: The Initial Value Problem*. John Wiley & Sons, 1991.

[12] J. C. Butcher, “Implicit Runge-Kutta Processes,” *Math. Comput.*, vol. 18, no. 85, pp. 50–64, 1964, doi: 10.2307/2003405.

[13] E. Hairer and H. Wanner, *Solving Ordinary Differential Equations II – Stiff and Differential-Algebraic Problems*, 2nd edition. Springer, 1996.

[14] C.-W. Shu, “Essentially Non-Oscillatory and Weighted Essentially Non-Oscillatory Schemes for Hyperbolic Conservation Laws,” NASA ICASE Report No. 97-65, Nov. 1997. Accessed: Jun. 06, 2020. [Online]. Available: https://ntrs.nasa.gov/search.jsp?R=19980007543

[15] L. Wang and D. J. Mavriplis, “Implicit solution of the unsteady Euler equations for high-order accurate discontinuous Galerkin discretizations,” *J. Comput. Phys.*, vol. 225, no. 2, pp. 1994–2015, Aug. 2007, doi: 10.1016/j.jcp.2007.03.002.

[16] A. Jameson, “Time dependent calculations using multigrid, with applications to unsteady flows past airfoils and wings,” in *10th Computational Fluid Dynamics Conference*, 0 vols., American Institute of Aeronautics and Astronautics, 1991. doi: 10.2514/6.1991-1596.

[17] B. C. Vermeire, N. A. Loppi, and P. E. Vincent, “Optimal Runge–Kutta schemes for pseudo time-stepping with high-order unstructured methods,” *J. Comput. Phys.*, vol. 383, pp. 55–71, Apr. 2019, doi: 10.1016/j.jcp.2019.01.003.

[18] J. D. Lambert, *Numerical Methods for Ordinary Differential Systems: The Initial Value Problem*. Wiley, 1991. [Online]. Available: https://books.google.co.kr/books?id=P0vPnQEACAAJ

[19] H,. T. Huynh, "Discontinuous Galerkin and Related Methods for ODE", Eleventh International Conference on Computational Fluid Dynamics (ICCFD11), Maui, Hawaii, USA, July 11-15, 2022

[20] M. Bogstad, D. Mavriplis and A. Kirby, "RANS and Hybrid RANS-LES Results for the Fourth High-Lift Prediction Workshop using the NSU3D Solver", AIAA Paper 2022-3810, AIAA Aviation 2022 Conference, Chicago IL, June 2022. https://doi.org/10.2514/6.2022-3810