Overview of the NASA Glenn Flux Reconstruction Based High-Order Unstructured Grid Code

Seth C. Spiegel,* James R. DeBonis† and H. T. Huynh,‡

NASA Glenn Research Center, Cleveland, OH, 44135, USA

A computational fluid dynamics code based on the flux reconstruction (FR) method is currently being developed at NASA Glenn Research Center to ultimately provide a large-eddy simulation capability that is both accurate and efficient for complex aeropropulsion flows. The FR approach offers a simple and efficient method that is easy to implement and accurate to an arbitrary order on common grid cell geometries. The governing compressible Navier-Stokes equations are discretized in time using various explicit Runge-Kutta schemes, with the default being the 3-stage/3\textsuperscript{rd}-order strong stability preserving scheme. The code is written in modern Fortran (i.e., Fortran 2008) and parallelization is attained through MPI for execution on distributed-memory high-performance computing systems. An \( h \)-refinement study of the isentropic Euler vortex problem is able to empirically demonstrate the capability of the FR method to achieve super-accuracy for inviscid flows. Additionally, the code is applied to the Taylor-Green vortex problem, performing numerous implicit large-eddy simulations across a range of grid resolutions and solution orders. The solution found by a pseudo-spectral code is commonly used as a reference solution to this problem, and the FR code is able to reproduce this solution using approximately the same grid resolution. Finally, an examination of the code’s performance demonstrates good parallel scaling, as well as an implementation of the FR method with a computational cost/degree-of-freedom/time-step that is essentially independent of the solution order of accuracy for structured geometries.

I. Introduction

Accurate and efficient numerical methods for the simulation of complex aerodynamic flows are desired to help reduce the design and development costs of aerodynamic vehicles and aeropropulsion systems. The primary challenge lies in the prediction of the turbulent motion within the flow and its effect on the overall motion of the fluid. Advanced methods such as large-eddy simulation (LES) and direct numerical simulation (DNS) offer great advancements in predicting turbulence, but these methods require significant improvements in accuracy and efficiency in order to become widely used. Focusing on either the accuracy or efficiency aspect of the underlying numerical algorithms, when developing a computational fluid dynamics (CFD) code, usually demands concessions from the other parameter.

Promising results have been found applying LES in the realm of aeropropulsion, but the spatial and temporal resolution needed for LES requires high-order (higher than second) numerical schemes historically only found in structured grid methods with large stencils. These methods have numerous drawbacks: difficulty in grid generation, high sensitivity to grid quality, complex boundary conditions, and poor scalability for parallel computing. These drawbacks severely limit the ability of these codes to compute complex aeropropulsion configurations, e.g., noise suppressing nozzles.\textsuperscript{1}

In an attempt to alleviate the difficulties of structured grid generation for complex geometries, researchers have applied LES methods to unstructured grids.\textsuperscript{2,3} However, the limited second-order accuracy of current unstructured methods requires an exorbitant number of grid points and offers little incentive over the traditional structured methods.\textsuperscript{4,5} The flux reconstruction (FR) method\textsuperscript{6–8} provides a simple, efficient, and
easy to implement method for solving the Navier-Stokes equations on unstructured grids and is accurate to an arbitrary order. By employing the high-order FR method with LES on unstructured grids, limitations related to structured grid generation, grid resolution, and computational efficiency can be overcome.9–12

Section II begins by providing an overview of the FR code that is the subject of this work. This overview includes a brief review of the governing equations in section II.A, brief descriptions of the numerical methods used to discretize the governing equations in space and time in section II.B, and a few details regarding the code development in sections II.C and II.D. Two canonical problems are then used in section III to verify the capabilities of this new code. The first problem consists of modeling a stationary isentropic vortex for an indefinite period of time, and the results are presented in section III.A. This version of the isentropic vortex is unchanging with time; therefore, an h-refinement study allows for a quantifiable measurement of the code’s order of accuracy since any deviation in the solution from the initial conditions is from numerical error. In section III.B, the Taylor-Green vortex problem is used to verify the implicit LES capabilities of this code for varying combinations of grid resolution and solution order. The results of these simulations are compared to those found by a pseudo-spectral code commonly used as a reference solution to this problem. Finally, a summary of the code and the numerical results are provided in section IV.

II. Code Description

A new CFD code named GFR (Glenn FR) is currently being developed at NASA Glenn Research Center to ultimately provide efficient LES capabilities using the FR method. This section describes some of the features and capabilities currently found within this new code, and begins with a brief review of the differential form of the compressible Navier-Stokes equations solved by GFR. Brief overviews of the FR method for discretizing the advection and diffusion terms of the governing equations are presented for the analogous 1D advection and diffusion equations, followed by the current methods available in GFR for advancing the solution in time. Finally, some details regarding the code development that separate the FR method (and related methods) from traditional finite-volume methods are discussed in the remaining sections.

II.A. Governing Equations

The 3D Navier-Stokes equations for a calorically perfect, compressible fluid can be written in differential form as

\[
\frac{\partial}{\partial t} (\rho) + \frac{\partial}{\partial x_i} (\rho v_i) = 0 \tag{1a}
\]

\[
\frac{\partial}{\partial t} (\rho v_i) + \frac{\partial}{\partial x_j} (\rho v_i v_j + \delta_{ij} p) = \frac{\partial}{\partial x_j} (\tau_{ij}) \tag{1b}
\]

\[
\frac{\partial}{\partial t} (\rho e_t) + \frac{\partial}{\partial x_i} (\rho v_i e_t + v_i p) = \frac{\partial}{\partial x_i} (v_j \tau_{ij} - q_i) \tag{1c}
\]

where \(\rho, v_i, p,\) and \(e_t\) respectively refer to the density, velocity in coordinate direction \(i\), pressure, and the total energy per unit mass of the system, and \(\delta_{ij}\) is the Kronecker delta function. Thermodynamic closure is found through the total energy equation

\[
\rho e_t = \frac{p}{\gamma - 1} + \frac{1}{2} \rho v_i^2 \tag{2}
\]

where \(\gamma\) is the ratio of specific heats for the fluid in question. The viscous stress tensor, \(\tau_{ij}\), is based on Stokes’ hypothesis and given by the equation

\[
\tau_{ij} = \mu \left( \frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3} \frac{\partial v_k}{\partial x_k} \delta_{ij} \right) \tag{3}
\]

Above, \(\mu\) is the molecular viscosity and its relation to the temperature of the fluid, \(T\), is modeled by Sutherland’s law

\[
\mu = \mu_0 \left( \frac{T}{T_0} \right)^{\frac{3}{2}} \frac{T_0 + T_{Suth}}{T + T_{Suth}} \tag{4}
\]
where $\mu_0$ is the reference viscosity at the reference temperature $T_0$ and $T_{Suth} = 110.4$ K. The heat flux vector, $q_j$, is modeled by Fourier’s law

$$q_j = -\lambda \frac{\partial T}{\partial x_j}$$

(5)

where $\lambda$ is the thermal conductivity which is related to the specific heat at constant temperature, $c_p$, and Prandtl number, $Pr$, through the relation

$$\lambda = \frac{c_p \mu}{Pr}$$

(6)

Additionally, with the specific gas constant denoted as $R_{\text{gas}}$, the temperature, pressure, and density of the fluid are related through the ideal gas law

$$p = \rho R_{\text{gas}} T$$

(7)

and the speed of sound, $a$, is given by the relations

$$a = \sqrt{\frac{\gamma p}{\rho}} = \sqrt{\gamma R_{\text{gas}} T}$$

(8)

In the present work, the fluid is assumed to be calorically perfect air with the following quantities held constant: the Prandtl number is $Pr = 0.72$, the ratio of specific heats is $\gamma = 1.4$, and the dimensionalized specific gas constant is $R_{\text{gas}} = 287.15$ J/(kg K).

II.B. Numerical Methods

The goal of this section is to present the base concepts behind the FR method and how it is implemented within GFR. We first summarize the discretization of the advection terms in equation (1) using the FR method initially presented by Huynh in Ref. 6. Next, the extension of FR for the diffusion terms is briefly discussed. Finally, we outline the methods for advancing the solution in time that are currently implemented within GFR.

The concepts of the FR method presented in this section are based on solving the 1D differential equations analogous to the advection and diffusion terms of the governing equations. This simplification is intended to give the reader a feel for how the method works, instead of a thorough and complex derivation of the method for a multidimensional system of coupled equations, e.g., the 3D Navier-Stokes equations. Tensor products are used to extend these ideas to higher dimensions for structured geometries, i.e., quadrilaterals and hexahedra, and detailed descriptions of the FR method for these geometries can be found in Ref. 6 and 7. Furthermore, details for extending the FR method to unstructured geometries can be found in Ref. 8 and 13, and Ref. 14 presents a matrix-oriented reformulation of the FR method that is generalized for all geometries.

II.B.1. Spatial Discretization of Advection Terms

To illustrate the idea of the FR method for the advection terms of equation (1), consider the 1D conservation law

$$u_t + f_x = 0$$

(9)

where the subscripts $t$ and $x$ denote partial differentiation with respect to time and space, respectively. Let the computational domain $\Omega$ be divided into an arbitrary number of nonoverlapping cells, and a mapping function, $r = \chi(x)$, is used to transform each cell from physical space to a common reference space/element. With $r$ denoting the local coordinate variable that varies on the reference element $I = [-1, 1]$, the inverse function $x = \chi^{-1}(r)$ similarly maps the reference element back to physical space. The details of such mappings, especially for higher dimensions and curved boundaries, are beyond the scope of this paper; the interested reader can find additional information in many texts on spectral-/finite-element type methods, including but not limited to Ref. 15–18. The remainder of this section focuses on using the FR method to solve the transformed version of equation (9) within each reference element.

At a given time $t$, the exact solution to equation (9) is approximated locally within each cell as a polynomial of degree $P$, and this polynomial approximation is generally discontinuous across cell interfaces. Let $\xi$, with $\xi_n$, $n = 1, 2, \ldots, N_P$, denote a set of coordinate values on $I$ that give the location of the $N_P = P + 1$ solution points within the reference element; the Gauss points are used as the solution points.
for all simulations in this work. Let \( u_n \) denote the known value of the solution polynomial at each solution point, i.e., \( u_n = u(\xi_n) \). The Lagrange polynomial is defined as the interpolating polynomial through a set of nodal points, for example \( \xi \), that takes the value 1 at \( \xi_n \) and zero at all other points \( \xi_i \neq n \):

\[
\ell_n(r) = \prod_{i: i \neq n}^{N_p} \frac{r - \xi_i}{\xi_n - \xi_i}
\]

(10)

The Lagrange polynomials serve as basis functions to construct the polynomial solution in each cell using

\[
u(r) = \sum_{n=1}^{N_p} u_n \ell_n(r)
\]

(11)

Since each \( u_n \) is simply a coefficient in equation (11), the derivative of this polynomial is just a linear combination of the derivatives of the Lagrange polynomials

\[
\frac{d}{dr} u(r) = u_r(r) = \sum_{n=1}^{N_p} u_n \left[ \frac{d}{dr} \ell_n(r) \right]
\]

(12)

Here and in the remainder of this work, the subscript \( r \) is exclusively used to denote differentiation with respect to the local coordinate variable \( r \) of a polynomial function defined on the reference element. Note that this same idea can be used to compute the derivative for any polynomial function whose values are known at the solution points.

Using the nodal solutions, we can evaluate the value of the flux function at each solution point, i.e., \( f^D_n = f(u_n) \). The superscript ‘D’ is used to identify it as the discontinuous flux function because the resulting function only uses information inside a given cell and does not account for interaction with data from the neighboring cells. Since we know the value of the discontinuous flux at each solution point in a cell, we can compute its derivative using equation (12). However, if we employ the derivative of \( f^D_n \) to evaluate \( f^D_r \) in equation (9), we obtain erroneous solutions since such a derivative does not include the interaction of data across cell interfaces. Therefore, we seek to construct a so-called continuous flux function, denoted by \( F(r) \), which accounts for this interaction between adjacent cells and approximates the discontinuous flux function in some sense, to which we can then calculate its derivative. The continuous flux function will be obtained by adding a correction to \( f^D_r \), the discontinuous one.

To this end, at each (cell) interface, let \( u_r \) and \( u_s \) be the values taken on by the solution polynomials to its left and right, respectively. An upwind flux value common for the two adjacent cells, denoted by \( f^* = f^*(u_r, u_s) \), can then be defined via the wind direction. Here, for the case of the Euler or Navier-Stokes equations, we use Roe’s flux\textsuperscript{19} with an entropy fix\textsuperscript{20}.

Let \( f^D_{\text{lb}} \) denote the upwind flux at the ‘left boundary’ of the cell and \( f^D_{\text{rb}} \) denote the upwind flux at the ‘right boundary’ of the cell. We can then construct the continuous flux function, \( F(r) \), as a polynomial of degree \( P + 1 \) by requiring that it takes on the upwind flux values at the two cell interfaces, and it approximates \( f^D \) via \( P - 1 \) conditions. Instead of defining \( F \), we define \( F - f^D \), which approximates the zero function. At the two interfaces, \( r = \pm 1 \), the function \( F(r) - f^D(r) \) takes on the following values, which are called the flux jumps:

\[
F(-1) - f^D(-1) = f^D_{\text{lb}} - f^D(-1) \quad \text{and} \quad F(1) - f^D(1) = f^D_{\text{rb}} - f^D(1)
\]

(13)

Therefore, \( F(r) - f^D(r) \) has prescribed left and right interface values, is of degree \( P + 1 \), and approximates zero.

We now separate the prescription of the jump at the left interface from that of the right. This separation plays a critical role in the new approach. Again, using the local coordinate, let \( g_{\text{lb}} \) be the correction function for the left interface of the reference element defined by

\[
g_{\text{lb}}(-1) = 1, \quad g_{\text{lb}}(1) = 0
\]

(14)

and \( g_{\text{lb}} \) is a polynomial of degree \( P + 1 \) approximating the zero function in some sense. We always choose \( g_{\text{rb}} \) by reflection: \( g_{\text{rb}}(r) = g_{\text{lb}}(-r) \). As a result,

\[
g_{\text{rb}}(-1) = 0, \quad g_{\text{rb}}(1) = 1
\]

(15)
All results within the present work use the Radau polynomials as the correction functions and this choice has been shown to recover a nodal discontinuous Galerkin (DG) method; further discussions regarding the choices in correction functions can be found in Ref. 6,21–23.

To begin constructing the continuous flux function, we start by considering the left interface, \( r = -1 \). The polynomial
\[
F(r) = f^p(r) + [f^p_{lB} - f^p(-1)] g_{lB}(r)
\]
provides a correction at the left interface for \( f^p(r) \) by changing the flux value at this interface from \( f^p(-1) \) to \( f^p_{lB} \), while leaving the value at the right interface unchanged, namely, \( f^p(1) \). Next, the polynomial
\[
F(r) = f^p(r) + [f^p_{lB} - f^p(-1)] g_{lB}(r) + [f^p_{RB} - f^p(1)] g_{RB}(r)
\]
provides corrections to both interfaces; using equations (14) and (15), one can easily verify that \( F(-1) = f^p_{lB} \) and \( F(1) = f^p_{RB} \). Thus, the above \( F(r) \) is of degree \( P + 1 \), takes on the upwind flux values at the two interfaces, and approximates \( f^p(r) \) in the same sense that \( g_{lB} \) and \( g_{RB} \) approximate the zero function.

The derivative of \( F(r) \) at the solution point \( \xi_n \) is evaluated as
\[
\frac{\partial F(r)}{\partial r}\bigg|_{\xi_n} = F_r(\xi_n) = f^p_r(\xi_n) + [f^p_{lB} - f^p(-1)] g^r_{lB}(\xi_n) + [f^p_{RB} - f^p(1)] g^r_{RB}(\xi_n)
\]
Finally, \( F_r \) is transformed from reference space to physical space using the inverse mapping function, \( \chi^{-1} \), where it can then be employed to approximate \( f_x \) at each solution point within the physical grid cell. The solution to equation (9) is then advanced in time via an explicit Runge-Kutta method.

II.B.2. Spatial Discretization of Diffusion Terms

We now move our attention to the diffusion terms in equation (1). Consider the 1D diffusion equation
\[
u t = u_{xx}
\]
which we choose to rewrite as a set of two coupled equations
\[
u u = \phi_x
\]
\[
\phi = u_x
\]
To solve equation (20a) and thus equation (19), we must first solve equation (20b) such that \( \phi \) is a polynomial of degree \( P \), approximates \( u_x \) within each cell, and takes into account the interaction of data between adjacent cells. This can all be accomplished using the reconstruction idea from the previous section. Switching back to the reference element, equation (18) can be reformulated to create an analogous expression for the reconstructed first derivative, \( \phi \), as
\[
\phi(\xi_n) = \phi_n = \phi^p_n + [u^*_{lB} - (u - 1)] g^r_{lB}(\xi_n) + [u^*_{RB} - u(1)] g^r_{RB}(\xi_n)
\]
In equation (21), the discontinuous first derivative, \( \phi^p_n = \phi^p(\xi_n) = u_r(\xi_n) \), is found using equation (12), and the common values at the interface, \( u^* \), are obtained by simply averaging the interface values from the two cells on the interface
\[
u u^* = \frac{1}{2} (u_l + u_r)
\]
Since \( \phi_n \) is now known at all solution points, a discontinuous second derivative, \( \phi^D_n \), can be evaluated at each solution point using the derivative of the Lagrange polynomials similar to equation (12), i.e.,
\[
\phi^D_r(r) = \sum_{n=1}^{N_P} \phi_n \left[ \frac{d}{dr} \ell_n(r) \right]
\]
Using the reconstructed first derivatives found by equation (21) and the subsequent discontinuous second derivatives, we can apply the reconstruction procedure one final time to evaluate the right side of equation (20a). The resulting equation for the corrected first derivative of \( \phi \), and thus the corrected second derivative of \( u \), is given by
\[
\frac{\partial \phi(r)}{\partial r}\bigg|_{\xi_n} = \phi^D_r(\xi_n) + [\phi^D_{lB} - \phi(-1)] g^r_{lB}(\xi_n) + [\phi^D_{RB} - \phi(1)] g^r_{RB}(\xi_n)
\]
All that remains is how to define the common derivatives, $\phi^*$, at each interface. Similar to equation (22), we use a simple average for the common derivative,

$$\phi^* = \frac{1}{2}(\phi_L + \phi_R) \quad (25)$$

where $\phi_L$ and $\phi_R$ are respective left and right derivative states on the interface. The obvious choice for the left and right derivative states is $\phi_L = \phi(1)$ for the left cell on the interface and $\phi_R = \phi(-1)$ for the right cell; however, this choice results in a large stencil with a relatively small stability limit. A compact stencil with a larger stability limit can be found through a one-sided correction. At each (cell) interface, let $\phi^0_L = u_r(1)$ and $\phi^0_R = u_r(-1)$ be the values taken on by equation (12) from the left and right cells on the interface, respectively. Application of the one-sided correction thus defines the left and right derivative states in the following way

$$\phi_L = \phi^0_L + (u^* - u_L)g'_{rb}(1) \quad (26a)$$
$$\phi_R = \phi^0_R + (u^* - u_R)g'_{lb}(-1) \quad (26b)$$

Using equations (22) and (26), and the identity $g'_{rb}(1) = -g'_{lb}(-1)$ resulting from the symmetry requirement on the correction functions, equation (25) can be simplified to

$$\phi^* = \frac{1}{2} \left[ \phi^0_L + \phi^0_R + (u_R - u_L)g'_{rb}(1) \right] \quad (27)$$

If the Radau polynomials are used for the correction functions in equations (21) to (27), the resulting method is equivalent to the BR2 method. All results in the present work use this method to evaluate the diffusion terms of equation (1).

II.B.3. Temporal Discretization

All of the methods currently implemented within GFR for advancing the solution in time are within the family of explicit Runge-Kutta (RK) schemes. The default scheme is the 3-stage strong stability preserving (SSP) RK method, which is 3rd-order accurate in time. Additional time integration methods available within GFR are the low-storage, 5-stage/4th-order RK method of Carpenter and Kennedy and the 5-stage/4th-order SSP RK method of Spiteri and Ruuth.

All of the results presented in this paper were integrated in time using the 3-stage SSP RK method. The time step size is evaluated at each interface flux point using the maximum inviscid and viscous eigenvalues from the two adjoining cells, and the minimum overall time step is used as the global time step to maintain a time-accurate, unsteady solution.

II.C. Data Structures and Code Optimization

The most efficient implementation of the FR method on structured grids utilizes the tensor product extension to higher dimensions. The tensor products can be programmed as a set of separate one-dimensional problems through an ordered, $ijk$-type data structure that is ideally suited for vectorization/optimization on modern computer processors. Unfortunately, such an implementation would severely handicap the applicability of a CFD code based on a method that is also known to work efficiently with unstructured geometries. The FR formulation for unstructured geometries is largely matrix-oriented such that many of the steps within the method can be simplified to matrix operations, e.g., matrix-vector products and dot products. A generalized method for all geometry types can be created by reformulating the tensor products of structured geometries in the same matrix-oriented formulation for unstructured geometries. Unlike the mostly full matrices for unstructured geometries, reformulating the tensor products results in very sparse matrices, thus eliminating the efficiencies inherent with the use of structured geometries. For 2D simulations, the sparsity of these matrices can be ignored without drastically affecting the performance of the algorithm. This cannot be said for 3D simulations where the zeros already comprise roughly 94% of these matrices for just a $P3$ solution. Combining this with a Courant-Friedrichs-Lewy (CFL) number that scales as $P^{-2}$ for explicit methods, running simulations above $P5$ becomes unreasonably expensive unless we can take advantage of the sparsity of these matrices.
Handling each geometry type separately from others would require the use of inefficient conditional statements at low levels within the solver, as well as multiple instances of duplicate code containing only small changes unique to each geometry. Therefore, one of our goals in developing GFR is to treat all grid geometries in a single, generic manner to enhance code clarity and readability. To facilitate this goal, a Fortran derived data type was created that stores a single matrix operator within an allocatable array for a specific combination of geometry type and solution order. For each matrix operator required by the FR method, we are able to create an array of this derived type that stores the matrices for each combination of geometry type and solution order valid for the current simulation.

With all of these matrix operators existing as an instance of this one derived data type, a generic mechanism for their use within the code was achieved through type-bound procedures, a new feature of Fortran that was introduced by the Fortran 2003 standard. By adding type-bound procedures with the pointer attribute to the definition of the derived type, the underlying algorithm controlling the matrix operations can be changed dynamically. Upon instantiation of each matrix object, the procedure pointers are associated with the Fortran intrinsic matrix functions (i.e., the \texttt{dot} product and \texttt{matmul} functions). After all the matrices are initially created during the preprocessor phase of the simulation, the sparsity of each matrix is evaluated. Each one that is found to be sparse is converted to a compressed storage format and stored within the derived type, and the procedure pointers are then associated with the sparse matrix algorithms that utilize this compressed format. Implementing these matrices in this object-oriented programming style creates additional flexibility in selecting which algorithms are utilized every time the matrix operators are used.

Since these matrix operations comprise a large percentage of the computations within each time step, this abstraction provides opportunities to significantly reduce the overall computational cost. As an example, our initial implementation of the sparse dot product turned out to be slower than the intrinsic version for a \( P3 \) simulation. Profiling of the sparse function highlighted significant overhead in evaluating the counter for the inner loop, because the compiler could not deduce that exactly four trips through the loop would occur every time the function was called. Therefore, the compiler was unable to properly optimize the sparse function by unrolling the inner loop, thus eliminating the need to evaluate the loop counter altogether. Using this information, we were able to quickly write new sparse functions with manually unrolled loops for many different solution orders, and associate these new functions with the type-bound procedure pointer.

Using the Taylor-Green vortex problem described in section III.B, we performed a series of tests to estimate the performance improvement found by the sparse algorithms. Each of the different matrix algorithms was used with \( P3-P8 \) solutions containing approximately \( 128^3 \) degrees of freedom (DoF), and each simulation was run for a total of 15,000 time steps. The total wall time for each simulation is shown in figure 1a and the speedup of the two sparse algorithms relative to the intrinsic functions is shown in figure 1b. The
most remarkable result from using the sparse algorithms is the computational cost per time step is relatively constant for all solution orders. In fact, the $\mathcal{P}5$ and $\mathcal{P}6$ solutions are actually less expensive per time step than the $\mathcal{P}3$ or $\mathcal{P}4$ solutions when comparing simulations with approximately the same DoF. Additionally, figure 1a clearly shows how expensive the solutions greater than $\mathcal{P}6$ become if the sparsity is ignored. It is also interesting to note that the intrinsic functions perform just as well as their sparse counterparts for the $\mathcal{P}3$ solution.

The matrix operators computed during the preprocessing phase are frequently used throughout each time step. Some specific examples of tasks involving these matrix operators include: computing derivatives through the derivative matrix, interpolating between sets of nodal points, use of the correction function, conversion between nodal and modal solution forms, and filtering. Roundoff errors occurring amid the creation of these operators are stored in the matrix elements; therefore, the errors are applied to the solution each and every time the operators are used. We take extra precaution to prevent this from happening by initially evaluating each operator using quadruple (128-bit) precision. Afterwards, each quad precision floating point value within each operator that is found to be smaller than double (64-bit) precision machine epsilon ($2^{-52}$) is assumed to be roundoff error and set to zero. All other values are simply converted from quad to double precision.

II.D. Parallel Processing

The parallel implementation is based on the single program, multiple data paradigm and utilizes MPI (Message Passing Interface)\textsuperscript{28} for communication between a set of processes on a distributed-memory system. The application programming interface (API) provided by the METIS\textsuperscript{29} library is used to partition the computational grid during the preprocessing phase of the simulation preventing the need to create any unnecessary intermediate grid files. Each MPI process corresponds to a single CPU core and the computational domain is partitioned such that each MPI process is assigned a unique grid partition. By default, the domain decomposition is accomplished through a $k$-way partitioning algorithm\textsuperscript{30} with the goal of minimizing the total communication volume required by the final partitioned grid. If the default method fails to produce a quality partitioning, i.e., each process/partition is assigned at least one grid cell and all partitions are contiguous, one of the other partitioning algorithms within METIS is attempted. The second method uses the same $k$-way partitioning algorithm, but the goal is changed to minimizing the edge cut\textsuperscript{31} which is just an approximation of the total communication required by the resulting partitioning. The final algorithm that is attempted uses a multilevel recursive bisection.\textsuperscript{29}

In cell-centered finite-volume (FV) methods, each grid cell contains only a single DoF which makes the computational cost per grid cell a small fraction of the total computational cost. However, for nodal high-order methods the DoF per grid cell is $O(d^d)$, where $d$ denotes the spatial dimension of the problem; therefore, the cost per grid cell is a much larger percentage of the total cost compared to FV methods. With communication significantly more expensive than computation, there is a limit for FV methods where the number of grid cells per partition becomes too small, and the performance starts to degrade rapidly as the communication costs begin to outweigh the computational costs. For nodal high-order methods, a partition can consist of just a few cells without severe communication bottlenecks, because the many DoF in each cell still require significant computational costs. The non-blocking communication functions in MPI are used to exchange information between processors in the background while computations local to each cell on a partition are evaluated concurrently to mask the communication costs. Exploiting these characteristics of high-order methods allows a problem to be partitioned efficiently across a very large number of processors with minimal communication costs.

Except for initially reading the global grid file, all large-scale file input/output (I/O) is done in parallel. There are two essential I/O tasks that inherently involve large amounts of data. The first task is the output of solution files used for visualizing the flow-field which is accomplished via the CGNS (CFD General Notation System) library.\textsuperscript{32} The CGNS library provides an API to create such files in a standardized, compact, machine-independent binary format that is compatible with most CFD flow analysis and visualization software packages. The popular HDF5 (Hierarchical Data Format, Version 5)\textsuperscript{33} data format is the default storage mechanism implemented within the CGNS library which allows for the utilization of the parallel I/O capabilities provided by the HDF5 libraries.

The second large-scale I/O task is the reading/writing of restart files. All restart file I/O within GFR is handled by the parallel I/O capabilities found in the MPI-2 extension of MPI.\textsuperscript{34} By utilizing MPI user-defined datatypes, all processes can collectively write their own local data to the restart file independent of all other
processes. The MPI datatypes automatically reorder the local data for each process such that it matches
the ordering of the global computational grid. Otherwise, the restart I/O would have to be performed either
serially (wasting significant memory and creating a desynchronized process) or by writing one or multiple
restart files that are dependent on the current partitioning.

The NASA Pleiades supercomputer was used to complete all of the simulations presented in this work. The
Pleiades supercomputer is a distributed-memory cluster consisting of SGI compute nodes, each based
on one of four different Intel CPU architectures, networked together via an InfiniBand interconnect and
connected to a Lustre parallel file system. All the simulations were run on either the “Sandy Bridge”-
based computing nodes that each contain two 8-core Intel Xeon E5-2670 processors or the “Haswell”-based
computing nodes that each contain two 12-core Intel Xeon E5-2680v3 processors. The largest simulation
attempted thus far used 4,800 CPU cores to successfully solve the Taylor-Green vortex problem on a grid
containing $102^3$ cells and 132,651,000 ($510^3$) DoF.

The most common method for testing parallel scalability, referred to as strong scaling, takes a problem
of fixed size and measures the speedup gained by increasing the number of processors in order to decrease
the amount of work assigned to each individual MPI process. A second test of parallel scalability, referred
to as weak scaling, counters strong scaling by scaling the problem size with the number of processors such
that the work assigned to each MPI process remains constant. Both of these methods were used to test
the parallel scalability of GFR through a series of $P_3$–$P_6$ simulations of the Taylor-Green vortex problem,
each solution order was run for 2,000 time steps on a successive number of Haswell compute nodes
($2^n$, $n = 0, 5$). Each Haswell node contains two 12-core processors, therefore the baseline simulation using
only a single compute node actually used 24 CPU cores; the remaining simulations used 48, 96, 192, 384,
and 768 CPU cores.

The strong scaling results for GFR are shown in figure 2a where the DoF for all simulations on all
processor counts was approximately $128^3$ DoF. Ideally, an increase in the number of processors should result
in a proportional reduction in the total wall time; this ideal speedup is shown as the black line in figure 2a.
Through 192 CPU cores, the strong scaling is within reason for all degrees of the solution polynomial, with
all showing a speedup within 92% of ideal at 192 cores. However, the scaling begins to digress with further
processor increases, especially so for the $P_5$ and $P_6$ simulations. At 768 CPU cores, the $P_5$ case contains
an average of only 12.1 cells per core and only 7.6 cells per core for the $P_6$ case. Because the total number
of cells cannot be split evenly between the grid partitions for these two cases, each MPI process contained

Figure 2: The parallel scaling found by $P_3$–$P_6$ simulations of the Taylor-Green vortex problem for 2,000
time steps on a varying number of CPU cores. The left figure shows the strong scaling for approximately
$128^3$ DoF split evenly across the CPU cores. The right figure shows the weak scaling with each CPU core
containing approximately 5,000 DoF.
either 12 or 13 cells for the \( P5 \) case or either 7 or 8 cells for the \( P6 \) case. These small one cell differences actually correspond to a significant load imbalance between processes, resulting in some waiting idly for others that have to do more work.

This inevitable limit of strong scaling is countered by weak scaling, and the results of the weak scaling tests with GFR are shown in figure 2b. The same successive number of compute nodes used for testing the strong scalability was also used to test the weak scalability, but here the simulations were scaled such that each CPU core contained approximately 5,000 DoF. The variations in the DoF in each simulation was accounted for by multiplying the overall wall time by the ratio 5,000/(actual DoF per CPU core). For all simulations using up to 16 computational nodes, the weak scaling is within 5% of the nominal scaling. However, the weak scalability begins to break down for all simulations using 32 nodes.

More than likely, both reductions in scalability at higher processor counts are caused by the increased complexity in the partition graph, thus creating significant contention for network bandwidth. At the moment, the domain decomposition and subsequent assignment of individual partitions to MPI processes is performed blindly such that the processor and partition localities are not reflective of one another. This method is more than adequate for lesser processor counts but it quickly breaks down as the number of processors increases; both are clearly reflected in figure 2.

III. Numerical Results

III.A. Isentropic Euler Vortex Problem

The isentropic Euler vortex problem is commonly used\(^9,13,16,35–46\) to empirically quantify the order of accuracy of a numerical method because it is easy to implement and the exact solution is known at all times. This problem is ideal for high-order methods because their theoretical accuracy should allow them to propagate the vortex with minimal entropy production for an indefinite amount of time. Being able to sustain vortical flow structures without the numerical method contributing unwanted numerical dissipation is crucial for advanced turbulence methods like LES. All of these reasons are why this is becoming a popular test case for high-order methods that are designed to produce minimal numerical dissipation.

In Ref. 45, some commonly used variations of this problem were found to contain instabilities caused by poorly chosen boundary conditions and the size of the computational domain being too small. After discovering the underlying causes of these instabilities, a list was provided to offer general guidelines and recommendations regarding the use of the isentropic Euler vortex problem. One of the recommendations is to zero the mean flow velocity so the vortex remains fixed at its initial location and no longer propagates through the domain. By keeping the vortex stationary, all of the grid boundaries can be given characteristic boundary conditions since periodic boundary conditions are no longer required. In the absence of physical dissipation, periodic boundaries provided a means for numerical error to recirculate within the computational domain and accumulate until eventually becoming large enough to cause the system to become unstable. The purpose of this article is to provide an overview of the capabilities of GFR, so this section will only summarize the findings from Ref. 45 regarding the stationary vortex problem with characteristic boundary conditions.

The initial conditions for the vortex problem are found by superimposing perturbations in velocity and temperature onto a uniform mean flow. These perturbations are defined such that the entropy, \( S = p/\rho^\gamma \), is constant throughout the entire grid domain, i.e., \( \delta S = 0 \). Since the vortex is stationary, the mean flow velocity, \((v_x,\infty, v_y,\infty)\), is zero; additionally, the free-stream density, \(\rho_\infty\), and pressure, \(p_\infty\), are set to unity.

The perturbations in velocity and temperature are defined as

\[
\begin{align*}
\delta v_x &= -y \Theta \\
\delta v_y &= +x \Theta \\
\delta T &= -\left(\frac{\gamma - 1}{2}\right) \Theta^2
\end{align*}
\] (28)

where \( \Theta \) is a Gaussian function of the form

\[
\Theta = \frac{5}{2\pi^{\gamma/2}} e^{-\frac{1}{2} (1-x^2-y^2)}
\] (29)
These perturbations, along with the free-stream conditions and the isentropic relations between density, pressure, and temperature, are used to define the initial flow conditions of the primitive variables as

\[
\hat{\rho}_0 = (1 + \delta T)^{\frac{1}{\gamma - 1}} \\
\hat{v}_{x,0} = \delta v_x \\
\hat{v}_{y,0} = \delta v_y \\
\hat{p}_0 = \frac{1}{\gamma} \left(1 + \delta T\right)^{\frac{\gamma}{\gamma - 1}}
\]

where the subscript 0 denotes a quantity given at \( t = 0 \). The tilde accents on the primitive variables indicate that each is a nondimensional quantity where \( \rho_\infty, a_\infty, \) and \( T_\infty \) have been used as the characteristic density, velocity, and temperature scales, respectively.

The last remaining item needed to complete the problem is the definition of the computational domain. All simulations of this problem were performed on a square domain with grid boundaries located at \((x,y) = (\pm 10, \pm 10)\). A series of grids with varying resolutions was created from the computational domain, with each grid consisting of regular, nonoverlapping quadrilateral cells of uniform size. Finally, all grid boundaries were set to characteristic boundary conditions to allow for information to flow in/out of the domain as needed. A contour plot showing the initial density profile resulting from this problem definition is shown in figure 3.

This problem is cleverly defined such that it can be run indefinitely and the exact analytical solution at any point in time is simply the initial conditions. Therefore, any deviation of the numerical solution from the initial conditions is an accurate measurement of the error generated by the numerical method, and a general sense of how the error evolves over time can be established. For the propagating vortex, time can additionally be measured in flow-through periods, which is the number of times the vortex propagates completely through the domain and returns back to its initial location. Even though a flow-through period for the stationary vortex is undefined, we will still use it to measure the elapsed time of our simulations. For the stationary case, one period is equivalent to the amount of time for a vortex propagating at \( v_{x,\infty} = 1 \) to complete one flow-through period.

The isentropic Euler vortex problem was used to complete an \( h \)-refinement study for \( P2 \rightarrow P4 \) solutions on grids containing \( 120^2 \), \( 160^2 \), and \( 200^2 \) cells. At each time interval that the solution error was evaluated, a linear fit of the density errors from each grid level was computed to see how the order of accuracy develops over time. The result of these linear fits using this set of simulations produces the order of accuracy versus time plot shown in figure 4. Initially, an accuracy of \( P+1 \) is found for each polynomial order, and this matches
the results commonly found by DG-type methods for solving nonlinear equations. As time progresses, the
accuracy increases for all polynomial orders reaching an approximate accuracy of $P + 3$ through 100 periods.
This clearly verifies that the FR method achieves some degree of super-accuracy, that is accuracy greater
than $P + 1$, for this problem.

III.B. Taylor-Green Vortex Problem

The Taylor-Green vortex (TGV) is a canonical problem in fluid dynamics developed to study vortex dynam-
ics, turbulent transition, turbulent decay and the energy dissipation process. The TGV problem contains
several key physical processes in turbulence in a simple construct and therefore is an excellent case for the
evaluation of turbulent flow simulation methodologies. The problem consists of a cubic volume of fluid
that contains a smooth initial distribution of vorticity, and periodic boundary conditions are applied to all
boundary surfaces. As time advances the vortices roll-up, stretch and interact, eventually breaking down into
turbulence. Because there is no external forcing, the small-scale turbulent motion will eventually dissipate
all the energy in the fluid and it will come to rest.

III.B.1. Problem Definition

The definition of the TGV problem used here follows that specified by the AIAA First International Workshop
on High-Order CFD Methods held in 2012. The same problem was also a test case in the 2nd and 3rd
iterations of workshop respectively held in 2013 and 2015, and is a test case for the upcoming 4th edition of
the workshop in 2016. The solution found by a pseudo-spectral code using 512 DoF is also included with
the workshop test case, and it is used here as a reference solution for comparison with the solutions found
by GFR.

The TGV problem models an incompressible flow at a Reynolds number of 1600 within a computational
domain defined by a periodic cube with sides of length $2\pi L$. Within the domain, an initial distribution of
velocity and its corresponding smooth vorticity distribution is specified by the following relations

$$
\begin{align*}
  v_x &= V_0 \sin \left( \frac{x}{L} \right) \cos \left( \frac{y}{L} \right) \cos \left( \frac{z}{L} \right) \\
  v_y &= -V_0 \cos \left( \frac{x}{L} \right) \sin \left( \frac{y}{L} \right) \cos \left( \frac{z}{L} \right) \\
  v_z &= 0 \\
  p &= p_0 + \frac{\rho_0 V_0^2}{16} \left( \cos \left( \frac{2x}{L} \right) + \cos \left( \frac{2y}{L} \right) \right) \left( \cos \left( \frac{2z}{L} \right) + 2 \right)
\end{align*}
$$

(31)

Because GFR solves the compressible Navier-Stokes equations, the reference Mach number is selected to
be small enough as to provide an incompressible flow at the given Reynolds number, i.e., $M_0 = 0.1$. To
complete the definition of the TGV problem, the reference length and temperature are respectively defined
as $L = 1.0 \text{ m}$ and $T_0 = 300 \text{ K}$.

III.B.2. Data Processing

The primary method for evaluating the TGV solutions is examining the rate at which the fluid dissipates
the kinetic energy in the domain. The integrated kinetic energy, $E_k$, within the domain is computed by

$$
E_k = \frac{1}{\rho_0 \Omega} \int_{\Omega} \rho \frac{v_i v_i}{2} \text{d}\Omega
$$

(32)

where $\Omega$ is the volume of the computational domain. The kinetic energy dissipation rate (KEDR) can then
be computed by differencing $E_k$ in time,

$$
\epsilon(E_k) = -\frac{dE_k}{dt}
$$

(33)

This KEDR, $\epsilon$, is written as a function of $E_k$ to emphasize that it is computed directly from the kinetic
energy in the domain.
A second “theoretical” kinetic energy dissipation rate can be computed from the integrated enstrophy, \( \zeta \), within the domain

\[
\zeta = \frac{1}{\rho_0 \Omega} \int_\Omega \rho \frac{\omega_i \omega_j}{2} d\Omega
\]  

(34)

where \( \omega_i \) is the vorticity vector. It can be shown that for an incompressible flow, the enstrophy is directly related to the kinetic energy dissipation rate through a constant,

\[
\epsilon(\zeta) = 2 \frac{\mu}{\rho_0} \zeta
\]  

(35)

The difference between the directly computed KEDR, \( \epsilon(E_k) \), and the enstrophy based KEDR, \( \epsilon(\zeta) \), will be examined.

Two additional quantities were computed to test the incompressible assumption. For a compressible flow where the bulk viscosity can be taken to be zero, the theoretical dissipation rate is the sum of two quantities; the contribution to the KEDR, obtained from the deviatoric strain rate tensor, \( S_{ij}^d \),

\[
\epsilon_1 = 2 \frac{\mu}{\rho_0 \Omega} \int_\Omega S_{ij}^d S_{ij}^d d\Omega
\]  

(36)

and the contribution to the KEDR, obtained from the pressure dilatation term,

\[
\epsilon_3 = -\frac{1}{\rho_0 \Omega} \int_\Omega \rho \frac{\partial v_i}{\partial x_j} d\Omega
\]  

(37)

For low Mach number flows, \( \epsilon_3 \) should be negligible and the theoretical dissipation rates from equations (35) and (36) should be equivalent. Even though GFR solves the compressible Navier-Stokes equations, a comparison between these two theoretical dissipation rates was nearly identical for each simulation, meaning the incompressible assumption is valid here.

III.B.3. Grids

Two input parameters are required to run the TGV problem using GFR. The first parameter simply instructs the code to initialize the flow solution using a function based on the relations in equation (31). The second input parameter is an integer which allows the user to load a user-provided grid if given a negative value. Otherwise, the parameter specifies the desired DoF in each coordinate direction, and a grid is generated during the preprocessor phase of the simulation based on this value and the input value for the degree of the polynomial solution. For example, if the input file requests 256 DoF and a \( P^4 \) solution (i.e., 5^3 DoF per grid cell), the code would generate a grid containing 51^3 cells resulting in a total of 255^3 DoF.

An optional input parameter can be given to apply a random perturbation to all of the internal grid nodes. For each individual grid node, a new random value is evaluated for the perturbation in each coordinate direction, so the highest possible degree of irregularity can be imposed onto the grid. The value of this parameter is input as a percentage, relative to the uniform cell length, and controls the maximum amount of perturbation allowed in any direction. The perturbations were limited to only the interior nodes of the grid to prevent the creation of hanging nodes across linked periodic boundaries or deformation to the cubic shape of the computational domain.

In this work, both \( h \)- and \( p \)-refinement studies were performed. For \( P^3 \)–\( P^6 \) solutions, simulations were performed on grids that provided approximately 64^3, 128^3, and 256^3 DoF for each of the solution orders. Additionally, simulations using \( P^7 \) and \( P^8 \) solutions were performed with approximately 256^3 DoF. Finally, a \( P^4 \) simulation was performed with approximately 512^3 DoF, and a second \( P^4 \) simulation with approximately 256^3 DoF was performed with a 20% random perturbation applied in all directions to each internal grid node. A summary of all these simulations is found in table 1 including details on the number of CPU cores used, the architecture of the CPU, and the average number of DoF and grid cells per CPU core.

III.B.4. Results

For the set of simulations with approximately 64^3 DoF, the \( P^5 \) and \( P^6 \) simulations blew up shortly after \( t^* = 15.04 \) and \( t^* = 10.44 \), respectively. This is clearly seen in figure 5a, and further attempts using successively smaller time steps failed to stabilize either of the simulations. A similar instability was found
using a $\mathcal{P}7$ solution with $128^3$ DoF, but the results of this simulation are not included since it only reached $t^* = 6.8$ before ultimately diverging. However, all of these instabilities were remedied by increasing the grid resolution through additional DoF as evident by the results in figure 7. This indicates that these instabilities could be driven by errors associated with polynomial aliasing, which are known to become more problematic as $\mathcal{P}$ increases for under-resolved simulations. Additional studies applying filtering or over-integration to these simulations are required to verify this conjecture. Also note that in figure 7, the $\mathcal{P}7$ and $\mathcal{P}8$ simulations do not reach $t^* = 20$, but this is not the result of instabilities causing the solution to diverge. It is simply because the simulations reached the end of their allocated wall time on the Pleiades system. The same is true for the $\mathcal{P}4$ simulation using $510^3$ DoF displayed in figure 8.

Figure 5, shows the KEDR for all the simulations containing approximately $64^3$ DoF; specifically, figure 5a shows the KEDR directly computed by time-differencing the integrated kinetic energy using equations (32) and (33), and figure 5b shows the KEDR based on the integrated enstrophy which is computed by equations (34) and (35). Aside from the points where the $\mathcal{P}5$ and $\mathcal{P}6$ solutions begin to diverge, the overall trends of $\epsilon(E_k)$ follow the spectral solution for all cases in figure 5a. However, in figure 5b, the KEDR is greatly under-predicted by $\epsilon(\zeta)$, especially around the time in which the spectral solution reaches a peak rate of dissipation, i.e., $t^* = 8.97$. For the $\mathcal{P}3$–$\mathcal{P}6$ solutions, the peak value is respectively under-predicted by 52%, 45%, 36%, and 28%, yet this also illustrates the considerable improvement in accuracy gained by increasing $\mathcal{P}$.

The directly computed KEDR is representative of all the dissipation actually present in the simulation, whereas the enstrophy based KEDR is representative of the physical dissipation that theoretically exists from the resolved vortical structures within the solution. Therefore, the difference between these two values can be considered as an approximation to the numerical dissipation provided by the FR method. In figure 5b, the physical dissipation from the resolved vortical structures is significantly less than the reference spectral solution, yet in figure 5a, the total dissipation follows the trends of the same reference solution. This indicates that the numerical dissipation is providing a reasonable approximation of the physical dissipation found in the unresolved scales. With this reasoning, the FR method appears to be a viable candidate for accurate implicit large-eddy simulations (ILES).

This same general trend is found with increasing grid resolution, as seen in figure 6 for simulations

<table>
<thead>
<tr>
<th>$\mathcal{P}$</th>
<th>DoF</th>
<th>Cells</th>
<th>CPU</th>
<th>Cores</th>
<th>Architecture</th>
<th>DoF/Core</th>
<th>Cells/Core</th>
<th>Error in $\epsilon(\zeta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>$64^3$</td>
<td>$16^3$</td>
<td>160</td>
<td>Sandy Bridge</td>
<td>1,638.4</td>
<td>25.6</td>
<td>51.58%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$65^3$</td>
<td>$13^3$</td>
<td>192</td>
<td>Haswell</td>
<td>1,430.3</td>
<td>11.4</td>
<td>45.40%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>$66^3$</td>
<td>$11^3$</td>
<td>192</td>
<td>Haswell</td>
<td>1,497.4</td>
<td>6.9</td>
<td>35.67%</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>$63^3$</td>
<td>$9^3$</td>
<td>120</td>
<td>Haswell</td>
<td>2,083.7</td>
<td>6.1</td>
<td>27.68%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\mathcal{P}$</th>
<th>DoF</th>
<th>Cells</th>
<th>CPU</th>
<th>Cores</th>
<th>Architecture</th>
<th>DoF/Core</th>
<th>Cells/Core</th>
<th>Error in $\epsilon(\zeta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>$128^3$</td>
<td>$32^3$</td>
<td>1,600</td>
<td>Sandy Bridge</td>
<td>1,310.7</td>
<td>20.5</td>
<td>24.16%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$130^3$</td>
<td>$26^3$</td>
<td>1,536</td>
<td>Haswell</td>
<td>1,430.3</td>
<td>11.4</td>
<td>19.26%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>$126^3$</td>
<td>$21^3$</td>
<td>720</td>
<td>Haswell</td>
<td>2,778.3</td>
<td>12.9</td>
<td>14.97%</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>$126^3$</td>
<td>$18^3$</td>
<td>480</td>
<td>Haswell</td>
<td>4,167.5</td>
<td>12.2</td>
<td>14.88%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\mathcal{P}$</th>
<th>DoF</th>
<th>Cells</th>
<th>CPU</th>
<th>Cores</th>
<th>Architecture</th>
<th>DoF/Core</th>
<th>Cells/Core</th>
<th>Error in $\epsilon(\zeta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>$256^3$</td>
<td>$64^3$</td>
<td>1,200</td>
<td>Haswell</td>
<td>13,981.0</td>
<td>218.5</td>
<td>5.96%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$255^3$</td>
<td>$51^3$</td>
<td>1,200</td>
<td>Haswell</td>
<td>13,817.8</td>
<td>110.5</td>
<td>3.69%</td>
<td></td>
</tr>
<tr>
<td>4*</td>
<td>$255^3$</td>
<td>$51^3$</td>
<td>2,400</td>
<td>Haswell</td>
<td>6,908.9</td>
<td>55.3</td>
<td>3.87%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>$258^3$</td>
<td>$43^3$</td>
<td>1,200</td>
<td>Haswell</td>
<td>14,311.3</td>
<td>66.3</td>
<td>3.04%</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>$259^3$</td>
<td>$37^3$</td>
<td>1,200</td>
<td>Haswell</td>
<td>14,478.3</td>
<td>42.2</td>
<td>2.08%</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>$256^3$</td>
<td>$32^3$</td>
<td>2,400</td>
<td>Haswell</td>
<td>6,990.5</td>
<td>13.7</td>
<td>2.01%</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>$252^3$</td>
<td>$28^3$</td>
<td>2,400</td>
<td>Haswell</td>
<td>6,667.9</td>
<td>9.1</td>
<td>1.40%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$510^3$</td>
<td>$102^3$</td>
<td>4,800</td>
<td>Haswell</td>
<td>27,635.6</td>
<td>221.1</td>
<td>0.75%</td>
<td></td>
</tr>
</tbody>
</table>

* Randomly perturbed grid
Figure 5: Comparison of the KEDR results for $P_3$–$P_6$ simulations containing approximately $64^3$ DoF.

Figure 6: Comparison of the KEDR results for $P_3$–$P_6$ simulations containing approximately $128^3$ DoF.
Figure 7: Comparison of the KEDR results for P3–P8 simulations containing approximately 256³ DoF.
Figure 8: Comparison of the KEDR results for the $P_4$ solutions on all of the various grid levels. The $P_8$ solution with $252^3$ DoF is also included in the close up views for additional comparisons.
containing approximately $128^3$ DoF, and again in figure 7 for those with approximately $256^3$ DoF. The last column of table 1 gives the error in $\epsilon(\zeta)$ for all of these simulations when the dissipation rate of the spectral solution is at its maximum. The error in $\epsilon(\zeta)$ decreases significantly with increasing grid resolution due to smaller and smaller vortical structures being resolved. This increase in the number of resolved vortical structures also corresponds to substantial improvements in $\epsilon(\mathcal{E}_h)$ converging towards the spectral solution. In figure 6a, the simulations using approximately $128^3$ DoF are much smoother and more accurate than those with $64^3$ DoF, and in figures 7a and 7c, the simulations containing approximately $256^3$ DoF are in very close agreement with the spectral solution.

The results in figures 5 to 7 present p-refinement studies of the KEDR at three different grid resolutions. The last set of KEDR results, presented in figure 8, demonstrates the effect of h-refinement by varying the DoF used with a P4 solution. The simulations containing $65^3$, $130^3$, and $255^3$ DoF are the same simulations from the previous figures. Included with these are a simulation containing $510^3$ DoF and a second simulation with $255^3$ DoF using the randomly perturbed grid described in section III.B.3. Close up views of figures 8a and 8b, in the region with peak KEDR, are respectively found in figures 8c and 8d. The P8 simulation with $252^3$ DoF has also been included in figures 8c and 8d to compare the improvement in accuracy found by increasing $P$ versus that found by increasing the grid resolution. Increasing the grid resolution to $510^3$ DoF (orange curve) gives slightly better accuracy in the KEDR over increasing the solution order to P8 (magenta curve), but both produce significant improvements in $\epsilon(\zeta)$ over the P4 solution with 2553 DoF (blue curve).

The next set of results for the Taylor-Green vortex problem demonstrate the accuracy of the actual vortical structures resolved by GFR. Contours of the vorticity norm at $t^* = 8$, located on the plane $x = -\frac{\pi L}{2}$ and in the region bounded by the lines $y = [0, \frac{\pi L}{2}]$ and $z = [\frac{\pi L}{2}, \pi L]$, are shown in figures 9 and 10 for each of the simulations with approximately $256^3$ DoF and in figure 11 for the P4 simulation with $510^3$ DoF. Additionally, the contour lines found using a pseudo-spectral code with 512$^3$ DoF have been superimposed onto each of these figures as a reference solution for comparison. Note that the contour lines correspond exactly to the contour color levels.

The solution from the spectral code resolves two primary vortical structures in the region of interest: a ring-like structure in the left half of the region and a linearly-shaped structure that extends from the bottom-right to just above the ring structure. In figure 9, both of these vortical structures from the spectral solution are resolved by GFR for all values of $P$. However, for lower orders, an additional structure not present in the spectral solution is found located at the bottom of this region. As the solution accuracy increases up to $P7$ and $P8$, this new structure is dissipated. One final point of interest in figure 9 is how the alignments of the finer details within the ring-like vortical structure change with increasing $P$. At lower orders, the interior edge of the ring structure is not well defined and the vorticity is spread across the interior of the ring. Additionally, the structure is more jagged and offset slightly in a counter-clockwise direction as compared to the spectral solution. As $P$ increases, the details of the ring structure converge towards the spectral solution with the $P8$ solution showing good agreement.

Figure 10 demonstrates the sensitivity of the FR method to irregularly spaced grids by comparing the two P4 simulations with $255^3$ DoF. The solution on the randomly perturbed grid in figure 10b is not significantly different from the solution on the uniform grid in figure 10a, but it does show increased diffusion throughout. The grid irregularity causes the interior edge of the ring structure to become even less defined due to additional spreading of the vorticity across the interior of the structure, and the edges of linear structure are smeared such that it deviates more from the spectral solution. The additional diffusion does provide some added benefit over the solution on the uniform grid, in that the structure not present in the spectral solution is dissipated and less prominent.

Finally, figure 11 shows that by increasing the grid resolution to $510^3$ DoF for a P4 solution, GFR is able to very nearly reproduce the spectral solution containing 512$^3$ DoF. The accuracy of this result provides some additional insight into the previous discussion using figures 8c and 8d to compare the accuracy improvement found by increasing grid resolution versus that found by increasing the solution order. Using the enstrophy based KEDR as the metric for accuracy, the simulation with additional grid resolution took longer than the simulation with an increased solution order, yet both gave similar improvements in accuracy. However, if the vorticity contours are used for the accuracy metric, comparing the improvement in accuracy going from figure 9b to 9f versus the improvement going from figure 9b to 11 shows that the increased grid resolution may be worth the extra computational cost. Combining all of the results for the TGV problem, the added value in both accuracy and flexibility provided by the two refinement methods, either separately or combined, is clearly visible.
Figure 9: Comparison of $P_3$–$P_8$ solutions with approximately $256^3$ DoF. The plots show contours of the vorticity norm at $t^* = 8$ on the plane $x = -\pi L$, in the region bounded by the lines $y = [0, \frac{\pi L}{2}]$ and $z = [\frac{\pi L}{2}, \pi L]$. The contour lines found using a pseudo-spectral code with $512^3$ DoF have been superimposed as a reference solution for comparison.
Figure 10: \( P^4 \) solutions with 255\(^3 \) DoF comparing the results between a uniform Cartesian grid and a randomly perturbed grid. The plots show contours of the vorticity norm at \( t^* = 8 \) on the plane \( x = -\pi L \), in the region bounded by the lines \( y = [0, \frac{\pi L}{2}] \) and \( z = [\frac{\pi L}{2}, \pi L] \). The contour lines found using a pseudo-spectral code with 512\(^3 \) DoF have been superimposed as a reference solution for comparison.

Figure 11: Contours from a \( P^4 \) simulation with 510\(^3 \) DoF showing the vorticity norm at \( t^* = 8 \) on the plane \( x = -\pi L \), in the region bounded by the lines \( y = [0, \frac{\pi L}{2}] \) and \( z = [\frac{\pi L}{2}, \pi L] \). The contour lines found using a pseudo-spectral code with 512\(^3 \) DoF have been superimposed as a reference solution for comparison.
Figure 12: Comparison of the aggregate wall times per DoF versus time step for each of the Taylor-Green vortex simulations. This shows that for each time step, the computational cost per DoF is essentially independent of $P$. Therefore, any additional computational cost incurred by increasing $P$ must be caused by the smaller time step restriction increasing the number of time steps required to reach a given simulation time ($t^*$).

III.B.5. Performance

All but two of the Taylor-Green vortex simulations were run on Haswell compute nodes; the $P^3$ solutions containing $64^3$ and $128^3$ DoF were run on Sandy Bridge nodes. Because these two CPU architectures prevent a fair performance comparison between simulations, a wall-time normalization factor was created using the TauBench benchmark program. Specifically, a series of five benchmarks was run on a single computational node of each type, making sure that: a) the same input parameters to TauBench were used for both node types, b) all CPU cores on the node were used for each benchmark, and c) the TauBench program was compiled for each node type using the same optimizations used to compile GFR. The averaged wall time for the five benchmarks on a given node architecture was then used to normalize the actual wall times for all simulations run on that same node architecture; the average of the benchmark wall times was 22.667 seconds for a Sandy Bridge node and 18.999 seconds for a Haswell node. Finally, since the number of CPU cores used for each simulation varied, the architecture-adjusted wall time was multiplied by the number of CPU cores to get an architecture-adjusted, aggregate wall time which provides a fair comparison of the computational cost between simulations.

For all the Taylor-Green vortex simulations, the computational cost per DoF is shown in figure 12 as a function of the number of completed time steps. There are two major deductions to be made from this figure. First, figure 12 simultaneously combines and reinforces the conclusions made from figures 1 and 2. Not counting the small benchmark cases from section II.C and section II.D, there are a total of 16 simulations of the Taylor-Green vortex problem presented in this work, and table 1 clearly shows the significant variations between all of the simulations in terms of grid size and the number of processors used. The fact that all 16 of these simulations are very closely aligned in figure 12 demonstrates a) the sparse matrix algorithms have essentially made the computational cost per time step independent of $P$, and b) the current parallel
implementation scales reasonably well for all combinations of $P$ and DoF used in this study. Second, if the cost per iteration is the same between two simulations with a similar number of DoF but different $P$, the difference in wall time to reach a given value of $t^*$ must be due to the difference in the number of time steps taken. Therefore, the stability restriction on the time step size for explicit methods remains one of the biggest obstacles preventing the use of FR, DG, and other high-order methods at very high orders, e.g., $P \geq 6$.

One last comment needs to be made about the results in figure 12 regarding the jumps in the wall time present in a number of the simulations. Each of these jumps coincides with a CGNS solution file being written to disk at predetermined simulation times of the TGV problem; specifically, the jumps occur when $t^*$ is 3, 5, 7, 8, 9, 11, and 15. At the time of this paper being prepared, the current stable release of CGNS, version 3.2.1, is the version implemented within GFR. This version of parallel CGNS, as well as all other previous versions, is known to experience major slowdowns when trying to open a file using a large number of processors. Version 3.3.0 is currently in its alpha release stage, and one of the major changes in this new update is significant performance improvements to the parallel CGNS functions.

IV. Summary and Conclusions

A new CFD code named GFR is currently being developed at NASA Glenn Research Center to ultimately provide efficient large-eddy simulation capabilities for complex aeropropulsion flows. A description of this code was presented, detailing the underlying numerical methods and some of its capabilities. This code uses the flux reconstruction method to solve the differential form of the Navier-Stokes equations, and various explicit Runge-Kutta schemes have been implemented for advancing the solution in time.

The FR method offers a simple and efficient method that is easy to implement and accurate to an arbitrary order on common grid cell geometries. A formulation of the FR method that unifies all grid geometries is achieved by reducing the method to a series of matrix operations. These matrix operators are very sparse for structured geometries, and type-bound procedure pointers, a new feature introduced by the Fortran 2003 standard, provide a path to significantly reduce computational costs by utilizing this sparsity. By implementing this strategy in GFR, the computational cost of using a $P$8 solution is nearly 5 times less than ignoring the sparsity and using the intrinsic functions.

Parallelization of GFR is achieved through MPI, and concurrently running computations and non-blocking communication are used to mask the expensive costs associated with communicating data between processors. Both weak and strong scaling tests show good results with minimal tuning of the parallel implementation. Our results indicate the scalability begins to decrease (to approximately 80–90% of ideal) for higher processor counts, but improvements are expected with adjustments to the partitioning algorithms and the method of distributing the partitions across the processors.

Using GFR, we were able to complete an $h$-refinement study of the isentropic Euler vortex problem and successfully verified the accuracy of the FR method for inviscid flows. Specifically, the results of the study found the expected $P + 1$ order accuracy through one period and super-accuracy, i.e., greater than $P + 1$, through 100 periods. The results from this case study demonstrate that vortical flow structures can be accurately predicted by the FR method for long amounts of time with minimal unwanted numerical dissipation introduced into the system.

The application of GFR to solve the Taylor-Green vortex problem demonstrated tremendous potential for its application to complex aerodynamic flows of engineering interest. The enstrophy based kinetic energy dissipation rate, $\epsilon(\zeta)$, is a theoretical measurement of the physical dissipation from the resolved vortical structures within the solution. This value was under-predicted in all simulations, especially for those containing more unresolved scales due to less grid resolution. However, the directly computed KEDR, $\epsilon(E_k)$, for each simulation followed the overall trend of the reference spectral solution with the highest grid resolutions showing very good agreement to this reference solution. The discrepancy between the theoretical and actual dissipation contained in the solution is representative of the numerical dissipation from the FR method. The accuracy of both $\epsilon(E_k)$ and $\epsilon(\zeta)$ consistently improved by increasing either or both grid resolution and $P$, which suggests that the numerical dissipation from the FR method adequately approximates the physical dissipation contained within the unresolved scales. Altogether, these results indicate that the FR method appears to be a viable candidate for accurate ILES of aerodynamic flows for complex geometries.
References


