Journal of Computational Physics 229 (2010) 6208–6219

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Toward textbook multigrid efficiency for fully implicit resistive magnetohydrodynamics

Mark F. Adams [a], Ravi Samtaney [b,*], Achi Brandt [c]

[a] Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, USA
[b] Princeton Plasma Physics Laboratory, Princeton University, Princeton, NJ, USA
[c] Department of Computer Science and Applied Mathematics, The Weizmann Institute of Science, Rehovot, Israel

ARTICLE INFO

Article history:
Received 29 October 2009
Received in revised form 5 April 2010
Accepted 12 April 2010
Available online 27 April 2010

Keywords:
Nonlinear multigrid
Defect correction
Implicit magnetohydrodynamics

ABSTRACT

Multigrid methods can solve some classes of elliptic and parabolic equations to accuracy below the truncation error with a work-cost equivalent to a few residual calculations – so-called "textbook" multigrid efficiency. We investigate methods to solve the system of equations that arise in time dependent magnetohydrodynamics (MHD) simulations with textbook multigrid efficiency. We apply multigrid techniques such as geometric interpolation, full approximate storage, Gauss–Seidel smoothers, and defect correction for fully implicit, nonlinear, second-order finite volume discretizations of MHD. We apply these methods to a standard resistive MHD benchmark problem, the GEM reconnection problem, and add a strong magnetic guide field, which is a critical characteristic of magnetically confined fusion plasmas. We show that our multigrid methods can achieve near textbook efficiency on fully implicit resistive MHD simulations.

1. Introduction

Resistive magnetohydrodynamics (MHD) is a fluid model for plasmas in which a single velocity and pressure field is used for both electrons and ions. Single-fluid resistive MHD is a system of nonlinear partial differential equations describing conservation of mass, momentum, and energy for an ionized gas coupled with Faraday's induction law, which describes the evolution of the magnetic field and Ohm's law. The resistive MHD model of a magnetized plasma does not include finite Larmor radius effects, and is based on the simplifying limit in which the particle collision length is small compared with the macroscopic length scales. For large magnetic Reynolds number or Lundquist number, the resistive MHD equations are dominated by the hyperbolic terms, except inside thin current layers whose thickness scales as the inverse square root of the Lundquist number. The hyperbolic portion of resistive MHD (aka ideal MHD) permits three speeds at which disturbances travel; these are the fast and slow magnetosonic speeds and the Alfvén speed. In a number of problems of physical interest, such as MHD investigations in a tokamak, there is a large temporal separation between the fast compressive motion and the dynamics of motion. In such situations, explicit time-integration methods are unnecessarily restrictive due to numerical stability constraints, whereas implicit methods are stable with large time steps. This paper develops fast, fully implicit nonlinear multigrid methods for eight field compressible single-fluid resistive MHD equations in conservation form.

Implicit and semi-implicit methods have been used in the context of MHD for many years in codes such as NIMROD (implicit) [23] and M3D (semi-implicit) [24], which were developed for tokamak MHD applications. Recent work in fully

* Corresponding author. Present address: Division of Physical Sciences and Engineering, KAUST, Thuwal 23955-6900, Saudi Arabia. Tel.: +1 966 2 808 2958.
E-mail address: samtaney@pppl.gov (R. Samtaney).

0021-9991/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2010.04.024

implicit MHD includes the application of Jacobian–Free–Newton–Krylov (JFNK) methods [7,21,20], which entails preconditioning a Krylov linear solver at each Newton iteration. Preconditioners are often important for effective convergence in JFNK, and linear multigrid can be used for this purpose [7]. Nonlinear multigrid – full approximation scheme (FAS) – is an alternative to JFNK, has been used for hyperbolic operators for many years by Brandt, Jameson and many others [5,19,11,16], and has the advantage being a globally convergent method. Furthermore, nonlinear multigrid obviates the need for an outer (Newton) iteration and can thus solve the nonlinear system directly with a cost similar to that of a linear solve. To the best of our knowledge nonlinear multigrid has never been applied to full eight field compressible resistive MHD, although there are a number of instances where these have been applied to hydrodynamics. Mavriplis [19] has used nonlinear multigrid successfully in the context of 2D unsteady flow. The work by Liao et al. [18] has achieved textbook efficiency on 3D unsteady compressible Navier–Stokes equations. They used a semi-Lagrangian approach for the advection part of the equations, while the viscous and pressure terms were discretized on a staggered mesh. A key development in their work is a distributed relaxation method to provide a stable smoother. Jameson and others use multi-stage Jacobi smoothers [9,16], which have the advantage that they are easier to implement given an existing time-stepping code and have fewer communications steps than Gauss–Seidel.

Minimizing the amount of work, or floating point operations (flop) per time step, is central to the design of fast solvers. Work complexity is a simple measure of computational cost that does not include, for instance, cache effects and other communication costs but is nonetheless very useful in understanding the potential performance of solvers. We define a "work unit" as the work equivalent to the cost of a residual calculation – this is the fundamental unit of work in fast iterative solvers. We count operations such as a Gauss–Seidel smoothing step, a first-order upwind operator, a second-order operator, and residual calculations as one work unit. The costs of each of these operations is not precisely the same but is close enough for our purposes. Moreover, we ignore the cost of interpolation and prolongation and simple vector operations. We consider this to be a simple and good complexity model for understanding the performance of these and other scalable solver methods. For an explicit multistage time-integration method the number of work units is the same as the number of stages. Common explicit methods might be second-order accurate and require two work units per time step. An implicit method with costs within a small factor of an explicit time step would be ideal for hyperbolic systems, and has been achieved with multigrid for some classes of parabolic problems [25]. This is known as "textbook" multigrid efficiency.

Multigrid is a powerful method whose ideas have been independently developed as, for instance, *nested iteration*, *error smoothing by relaxation*, and *total reduction*. Multigrid, or more generally multi-level, methods are the only theoretically optimal solvers with serial work complexity of $\mathcal{O}(N)$ and $\mathcal{O}(\log(N))$ parallel complexity for a V-cycle as defined in Section 3 (an F-cycle, as defined in Section 3, adds a $\log(N)$ term to each complexity), where N is the total number of grid points or cells. These are theoretically optimal limits on the complexity of solving an elliptic system of equations. Multigrid methods have been shown to be very efficient and practical for elliptic operators and a rich set of theories provides strong optimality bounds for elliptic operators [4,10,14,25]. On the standard model problem, the Laplacian, multigrid with linear interpolation/restriction, Gauss–Seidel smoothers, and F-cycles reduce the algebraic error to well below that of the discretization error in one iteration with about seven work units (see Section 3 for details). Achieving this textbook efficiency for operators of interest in science and engineering is an active area of research. The challenge for hyperbolic dominated nonlinear PDEs is the development of stable and efficient smoothers. To this end, we develop a nonlinear multigrid method for resistive MHD using stable pointwise Gauss–Seidel smoothers and standard geometric nonlinear multigrid. An additional consideration in algorithmic design and implementation is that recent machines have limited memory bandwidth relative to increasing processor flop rates, and this trend is expected to continue with future generations of machines. Most of the machine cycles are spent waiting for memory to be brought to the arithmetic units and most of the power budget is spent on moving data. Minimizing the use of memory is thus an important criterion in the design and implementation of equation solvers for future machines. The geometric multigrid methods used here are ideally suited for these new architectures because they can naturally be implemented without explicitly storing a matrix (i.e., *matrix-free*), and in fact these methods were developed in the seventies for much the same reasons when memory availability was limited.

This paper proceeds with an introduction to MHD equations and discretizations in Section 2, followed by a description of our multigrid method applied to MHD in Section 3, numerical results in Section 4, and conclusions and future work in Section 5.

2. Magnetohydrodynamics: equations and discretization

2.1. Resistive MHD equations

The single-fluid resistive MHD equations couple the equations of hydrodynamics and resistive Maxwell's equations, and may be written in conservation form as

$$\frac{\partial U}{\partial t} + \underbrace{\nabla \cdot F(U)}_{\text{hyperbolic terms}} = \underbrace{\nabla \cdot F_d(U)}_{\text{diffusive terms}}, \tag{1}$$

where the solution vector $U \equiv U(\mathbf{x},t)$ is

$$U = \{\rho, \rho\mathbf{u}, \mathbf{B}, e\}^T,$$

and the hyperbolic flux vector $F(U)$ and the diffusive fluxes $F_d(U)$ are given by

$$\begin{aligned} F(U) &= \left\{\rho\mathbf{u}, \rho\mathbf{u}\mathbf{u} + \left(p + \frac{1}{2}\mathbf{B}\cdot\mathbf{B}\right)\bar{\bar{\mathbf{I}}} - \mathbf{B}\mathbf{B}, \mathbf{u}\mathbf{B} - \mathbf{B}\mathbf{u}, \left(e + p + \frac{1}{2}\mathbf{B}\cdot\mathbf{B}\right)\mathbf{u} - \mathbf{B}(\mathbf{B}\cdot\mathbf{u})\right\}^T, \\ F_d(U) &= \left\{0, Re^{-1}\bar{\bar{\boldsymbol{\tau}}}, S^{-1}(\eta\nabla\mathbf{B} - \eta(\nabla\mathbf{B})^T), Re^{-1}\bar{\bar{\boldsymbol{\tau}}}\cdot\mathbf{u} + \frac{\gamma}{\gamma-1}\frac{\kappa}{RePr}\nabla T + \frac{\eta}{S}\left(\frac{1}{2}\nabla(\mathbf{B}\cdot\mathbf{B}) - \mathbf{B}(\nabla\mathbf{B})^T\right)\right\}^T. \end{aligned} \tag{2}$$

In the above equations ρ is the density, $\mathbf{u}$ is the velocity, $\mathbf{B}$ is the magnetic field, p and T are the pressure and temperature, respectively, and e is the total energy per unit volume of the plasma. The plasma properties are the resistivity η, the thermal conductivity κ, and the viscosity μ, which have been normalized, respectively, by a reference resistivity η_R, a reference conductivity κ_R, and a reference viscosity μ_R. The ratio of specific heats is denoted by γ and taken to be 5/3 throughout this work. The non-dimensional parameters in the above equations are the Reynolds number, defined as $Re \equiv \rho_0 U_0 L/\mu_R$, the Lundquist number, defined as $S \equiv \mu_0 U_0 L/\eta_R$, and the Prandtl number, denoted by Pr, which is the ratio of momentum to thermal diffusivity. The non-dimensionalization is carried out using a characteristic tokamak length scale, L, and the Alfvén speed $U_0 = B_0/\sqrt{\mu_0\rho_0}$, where B_0, ρ_0, and μ_0 are the characteristic strength of the magnetic field, a reference density, and the permeability of free space, respectively. The equations are closed by the following equation of state:

$$e = \frac{p}{\gamma-1} + \frac{\rho}{2}\mathbf{u}\cdot\mathbf{u} + \frac{1}{2}\mathbf{B}\cdot\mathbf{B},$$

and the stress tensor is related to the strain as

$$\bar{\bar{\boldsymbol{\tau}}} = \mu\left(\nabla\mathbf{u} + (\nabla\mathbf{u})^T\right) - \frac{2}{3}\mu\nabla\cdot\mathbf{u}\bar{\bar{\mathbf{I}}}.$$

Finally, a consequence of Faraday's law is that an initially divergence-free magnetic field must lead to a divergence-free magnetic field for all times, which corresponds to the lack of observations of magnetic monopoles in nature. This solenoidal property is expressed as $\nabla\cdot\mathbf{B} = 0$.

2.2. Discretization

We limit the discussion to a two-dimensional Cartesian domain. The extension to three dimensions is straight-forward. We use a finite volume approach in which the domain $[x_l,x_r]\times[y_l,y_r]$ is divided into finite volumes of size $\Delta x\times\Delta y$ (Δx and Δy are the mesh spacing in the x- and y-directions, respectively), indexed by (i,j) and bound by faces at $(i+\frac{1}{2},j)$ and $(i-\frac{1}{2},j)$ in the x-direction, and $(i,j+\frac{1}{2})$ and $(i,j+\frac{1}{2})$ in the y-direction. Solving Eq. (1) using a backward Euler method:

$$A\left(U_{ij}^{n+1}\right) \equiv U_{ij}^{n+1} + \Delta t\left[\nabla\cdot F\left(U^{n+1}\right) - \nabla\cdot F_d\left(U^{n+1}\right)\right]_{ij} = U_{ij}^{n}. \tag{3}$$

We discretize the divergence of the fluxes in Eq. (1) as

$$\left(\frac{\partial f}{\partial x}\right)_{i,j} = \frac{\tilde{f}_{i+\frac{1}{2}j} - \tilde{f}_{i-\frac{1}{2}j}}{\Delta x}, \tag{4}$$

where f may represent either the hyperbolic or the diffusive fluxes. The quantity $\tilde{f}_{i+\frac{1}{2}j}$ is referred to as the numerical flux through the face $\{i+\frac{1}{2},j\}$, and is computed as a linear combination of the fluxes at cell centers as

$$\tilde{f}_{i+\frac{1}{2}j} = \sum_{v=-m}^{n} a_v f_{i+v,j}. \tag{5}$$

Our numerical framework provides the flexibility of using any of a group of different spatial discretization schemes. For a second-order central difference implementation, $m = 0$, $n = 1$, and $a_0 = a_1 = \frac{1}{2}$; for a fourth-order central difference approximation, $m = 1$, $n = 2$, and $a_{-1} = a_2 = -\frac{1}{12}, a_0 = a_1 = \frac{7}{12}$. These central difference approximations are free of dissipation errors, except perhaps near domain boundaries. However, for physical phenomena which are not adequately resolved, the numerical solution will exhibit ringing due to the dispersion error in these central difference schemes. In the next section, we present a nonlinear multigrid method to solve the discrete system (Eq. 3), the left-hand side of which is denoted as $A(U_{ij}^{n+1})$, and which uses the central difference discretizations to evaluate the divergence of the fluxes. We utilize an additional discretization of Eq. (3) where the hyperbolic fluxes are computed using a first-order upwind method. This operator is denoted as $\widetilde{A}(U_{ij}^{n+1})$. The first-order upwind fluxes are computed using a Lax–Friedrichs method as follows:

$$F_{i+\frac{1}{2}j} = \frac{1}{2}\left[F(U_{ij}) + F(U_{i+1,j})\right] - \frac{1}{2}|\lambda_{\max}|(U_{i+1,j} - U_{ij}), \tag{6}$$

where $\lambda_{\max}$ is the maximum eigenvalue of the Jacobian matrix $\partial F/\partial U$, associated with the fast compressive wave in MHD.

3. Multigrid methods

Multigrid methods are motivated by the observation that a low resolution discretization of an operator can capture modes or components of the error that are expensive to compute directly on a highly resolved discretization. More generally, any poorly locally-determined solution component has the potential to be resolved with coarser representation. This process can be applied recursively with a series of coarse grids, thereby requiring that each grid resolve only the components of the error that it can solve efficiently. These coarse grids have fewer grid points, typically about a factor of two in each dimension, such that the total amount of work in multigrid iterations can be expressed as a geometric sum that converges to a small factor of the work on the finest mesh. These concepts can be applied to problems with particles/atoms or pixels as well as the traditional grid or cell variables considered here. Multigrid provides a basic framework within which particular multigrid methods can be developed for particular problems.

We use geometric multigrid because it has the potential to be very efficient and the geometric domains of interest in this work (i.e., tokamaks) are simple enough that explicit coarse grids can be practically constricted even if, for instance, unstructured grids are used. Geometric multigrid not only provides a powerful basis on which to build a specific solution algorithm, but also allows for the straightforward use of nonlinear multigrid, or full approximation scheme (FAS) multigrid [4] and matrix-free implementations. Given that our problems are nonlinear, FAS multigrid is very efficient in that it solves the nonlinear set of equation directly and obviates the need of an outer (Newton) iteration. This is a critical component in attaining textbook efficiency on nonlinear problems. Fig. 1 shows the standard multigrid FAS V-cycle and uses the smoother $u \leftarrow S(A,u,b)$, the restriction operator R_k^{k+1}, which maps residuals and current solutions from the fine grid space k to the coarse grid space $k+1$ (the rows of R_k^{k+1} are the discrete representation, on the fine grid, of the coarse grid functions), and the prolongation operator P_{k+1}^k, which maps the current solution from the coarse grid to the fine grid.

Common notation for this multigrid V-cycle is V (μ_1,μ_2), where μ_1 and μ_2 are the number of pre- and post-smoothing steps, respectively. The canonical model problem is the Laplacian, for which point-wise Gauss–Seidel smoothers combined with linear interpolation for the restriction and prolongation operators generate methods that reduce error by about an order of magnitude per V (1,1) cycle. This is theoretically optimal in that this rate of residual reduction is independent of mesh size and the amount of work in a V-cycle is given by a geometric sum that converges to about five work units. This so-called textbook efficiency has been observed, if not proven, for multigrid methods in a wide variety of applications (see Trottenberg et al. and references therein for details [25]).

A concept used to determine if a point-wise smoothing method exists is *h-ellipticity*. Brandt et al. in [6] first introduced *h*-ellipticity and it is described in Trottenberg et al. [25]. *H*-ellipticity is the minimum Fourier symbol of the high half (in at least one dimension) of the spectrum of a discrete operator divided by the maximum Fourier symbol of the operator. An *h*-ellipticity bounded well above zero is a necessary and sufficient condition for the existence of a point-wise smoother for an operator with a symmetric stencil [25]. An important result of *h*-ellipticity is that effective point-wise smoothers (e.g., Gauss–Seidel and distributive Gauss–Seidel) can be constructed for upwind discretizations of hyperbolic systems with no restriction on the time step, whereas point-wise Gauss–Seidel is unstable for a central difference scheme for a large time step. We have observed textbook multigrid efficiency with standard multigrid methods (e.g., point-wise Gauss–Seidel smoothers) using a first-order upwinding method for ideal and resistive MHD. First-order accuracy is, however, generally not sufficient for many applications, and second-order schemes are required for efficiency. These stable low-order smoothers have been used extensively with a higher-order operator via a defect correction scheme, which is identical to preconditioning when an exact solver is used, but is more amenable to a nonlinear solve [1,3,8,15,17].

An additional requirement of an optimal solver is to be able to reduce the algebraic error to the order of the discretization (or truncation) error for steady-state problems. For transient problems the solver needs to reduce the algebraic error to below the *incremental* error – that is, the product of the truncation error of the time-integration scheme and the spacial truncation error. Reducing algebraic error far below that of the incremental error is computationally wasteful, though potentially useful for debugging. There is generally no need to spend resources to reduce the algebraic error far below the incremental

```
function u ← MGV(A_k, u_k, f_k)
    if coarse grid k+1 exists
        u_k ← S(A_k, u_k, f_k)                    /* pre-smoothing */
        r_k ← f_k − A_k u_k
        r_{k+1} ← R_k^{k+1}(r_k)                  /* restriction of residual to coarse grid */
        w_{k+1} ← R_k^{k+1}(u_k)                  /* restriction of current solution to coarse grid */
        u_{k+1} ← MGV(A_{k+1}, w_{k+1}, r_{k+1} + A_{k+1} w_{k+1}) /* recursive multigrid application */
        u_{k+1} ← u_{k+1} − w_{k+1}               /* convert solution to an increment */
        u_k ← u_k + P_{k+1}^k(u_{k+1})            /* prolongation of coarse grid correction */
        u_k ← S(A_k, u_k, f_k)                    /* restriction of residual to coarse grid */
    else
        u_k ← A_k^{-1} f_k                        /* post-smoothing */
    return u_k
```

Fig. 1. Nonlinear FAS multigrid V-cycle algorithm.

```
function u ← MGF(A_k, u_k, f_k)
    if coarse grid k+1 exists
        r_k ← f_k − A_k u_k
        r_{k+1} ← R_k^{k+1}(r_k)                    /* restriction of residual to coarse grid */
        w_{k+1} ← R_k^{k+1}(u_k)                    /* restriction of solution to coarse grid */
        u_{k+1} ← MGF(A_{k+1}, w_{k+1}, r_{k+1} + A_{k+1} w_{k+1}) /* recursive multigrid application */
        u_{k+1} ← u_{k+1} − w_{k+1}                 /* convert solution to an increment */
        u_k ← u_k + P_{k+1}^k(u_{k+1})              /* prolongation of coarse grid correction */
        u_k ← MGV(Ã_k, u_k, f_k − A_k u_k + Ã_k u_k) /* low-order V-cycle, defect correction */
    else
        u_k ← A_k^{-1} f_k                          /* accurate solve of coarsest grid */
    return u_k
```

Fig. 2. Nonlinear FAS multigrid F-cycle algorithm with defect correction.

error. This observation leads to our definition of an optimal solver as one that can reduce the error to less than the incremental error with a few work units per time step. This is an ambitious goal in that it requires both scalability and small constants in the actual computational costs. In fact, this results in a solver in which the rate of reduction in the residual actually increases as the mesh is refined, because the truncation error decreases. This goal can be achieved by using a multigrid V-cycle within what is called an F-cycle iteration [25]. Fig. 2 shows the standard nonlinear multigrid F-cycle with defect correction to accommodate the nonlinear V-cycle with a lower-order operator ($\tilde{A}$ is the first-order upwinding operator) for which our point-wise Gauss–Seidel smoother is stable.

The complexity of an F-cycle is asymptotically similar to a V-cycle, and it can be proven to result in a solution with algebraic error that is less than the incremental error on the model problem [25]. Multigrid can thus achieve discretization error with a work complexity of a few residual calculations by using multigrid. An additional advantage of the FAS multigrid algorithm is that it is an effective global nonlinear solver in that it does not suffer from the problem of limited radius of convergence of a standard Newton method.

4. Numerical results

4.1. Test problem setup and solver definition

We evaluate the effectiveness of our methods with a resistive MHD benchmark problem in magnetic reconnection – the geospace environmental modeling (GEM) magnetic reconnection problem [2]. The GEM problem is a Harris sheet equilibrium with a magnetic field perturbation which, in the presence of resistivity, allows for the parallel magnetic field lines to break and reconnect.

Fig. 3 shows a time sequence of current density J_z field during reconnection.

Two Reynolds numbers are considered in Eq. (2): (1) $Re = 2000$ *low viscosity*, and (2) $Re = 20$ *high viscosity*. The Lundquist number was fixed at $S = 200$ and the thermal conduction term was chosen to be $Re\,Pr = 125$. The boundary conditions are periodic in x and perfectly conducting walls in y. Two magnetic field configurations are also considered: (1) zero B_z field and (2) $B_z = 5$ field. We consider large magnetic guide fields because this is an important parameter of magnetically confined fusion plasmas. A guide field make these problems much stiffer; for instance, in the limit of an infinite guide field, the velocity field becomes incompressible. The two Reynolds numbers and two guide field cases generate four test scenarios that we investigate.

Our goal is to develop solvers with a complexity equivalent to a few residual calculations (work units) per time step, with the largest time step that can accurately resolve the dynamics of the problem. In this study, the solver is fixed at one iteration of FAS F-cycle with two defect corrected V (1,1) cycles at each level, as described in Section 3, and with a work complexity of about 18 work units per time step. There are three applications of the fine grid operator in residual calculations and defect correction in FAS multigrid, and three fine grid work units in the smoothers and residual calculations in each of the two V (1,1) cycle, plus lower-order work in restriction/prolongation and FAS terms. This results in about ten work units on the fine

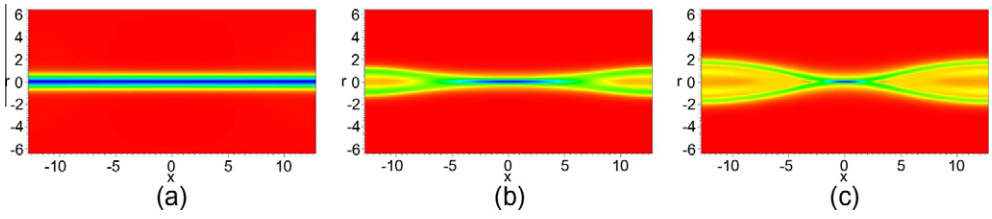

Fig. 3. Time sequence of current density, J_z during reconnection at time (a) $t = 0$, (b) $t = 15$ and (c) $t = 60$, for the low viscosity zero-guide field case.

grid. Each successive grid is four times smaller (in 2D), and F-cycles process the second grid twice, the third grid three times, and so on, resulting in the equivalent of about eight additional work units for a total of 18 work units (there are actually fewer total work units in 3D because the coarse grids are relatively smaller). The smoother is nonlinear Gauss–Seidel with one iteration per grid point and red–black (or checkerboard) ordering. In Section 4.2 we demonstrate, with comparisons against several MHD codes, that we are solving the equations correctly. In Section 4.3 we present convergence studies in space and time that demonstrate that we are, in fact, able to attain second-order spatial accuracy of our numerical method with this nonlinear multigrid MHD solver. The code is fully MPI parallel and we observe good parallel efficiency on typical current high-end parallel machines (Section 4.5). For the test cases presented here, the time integrator is backward Euler, which is first-order accurate. The code has the option of second-order Crank–Nicholson time stepping, but we observed problems with stability on the stiffer test cases. There are two paths to higher-order in time solutions: (1) develop higher-order stable time integrators or (2) use a higher-order time integrator in the defect correction process. The latter of these options is simple, and we plan to investigate this in future work. We use a time step $\Delta t = 0.1$ for all of the test cases, even though the guide field cases are much stiffer. This is about twice as expensive as the "textbook" efficiency that was our goal, and we are able to use an optimal one V(1,1) cycle for the high viscosity, zero-guide field case (as well as Crank–Nicholson), but this solver was not adequate on the other test problems with our time step, and so for simplicity of presentation we use the same solver parameters for all test cases.

4.2. Verification

To verify that we are solving the equations accurately we compare the evolution of the total kinetic energy with four other resistive MHD codes: (1) an explicit finite volume code (labeled "finite volume" [22]), (2) a C1 finite element code (labeled "M3D-C1" [12]), (3) a spectral element code (labeled "SEL" [13]), and (4) a higher-order finite element code (labeled "NIMROD" [23]). The comparison, using the high viscosity case with zero magnetic guide field, is shown in Fig. 4 (left).

Our code agrees with these other codes with a peak kinetic energy of about 0.22 at about $T = 9.0$ and minimum at $T = 30.0$.

Similarly, Fig. 5 (left) shows the low viscosity case. Again, good agreement is observed with all codes with a peak kinetic energy of about 0.46 at about $T = 14.0$ and 0.12 at $T = 40.0$. Due to the damping nature of backward Euler time integration, the low amplitude oscillations in the kinetic energy history, observed in the other codes, are damped. As expected, the low viscosity case requires a finer grid than the high viscosity case.

Finally, a high viscosity case with a magnetic guide field of $B_z = 5.0$ is compared with a C1 finite element [12] code in Fig. 6 (left). Again, good agreement is observed, although the finite element results are not highly converged.

4.3. Convergence tests

The ultimate goal of solver design is to be able to take as large a time step as possible while capturing the dynamics of the system accurately and achieving second-order spatial accuracy. To verify that our solver is indeed reducing the algebraic error below truncation error, as F-cycle multigrid theory would indicate, we rely on empirical evidence of second-order spatial convergence. Observing second-order spatial accuracy is sufficient to verify that the solver is truly optimal in that the algebraic error is reduced to the order of the discretization error with an asymptotically constant amount of work per grid point (i.e., 18 work units).

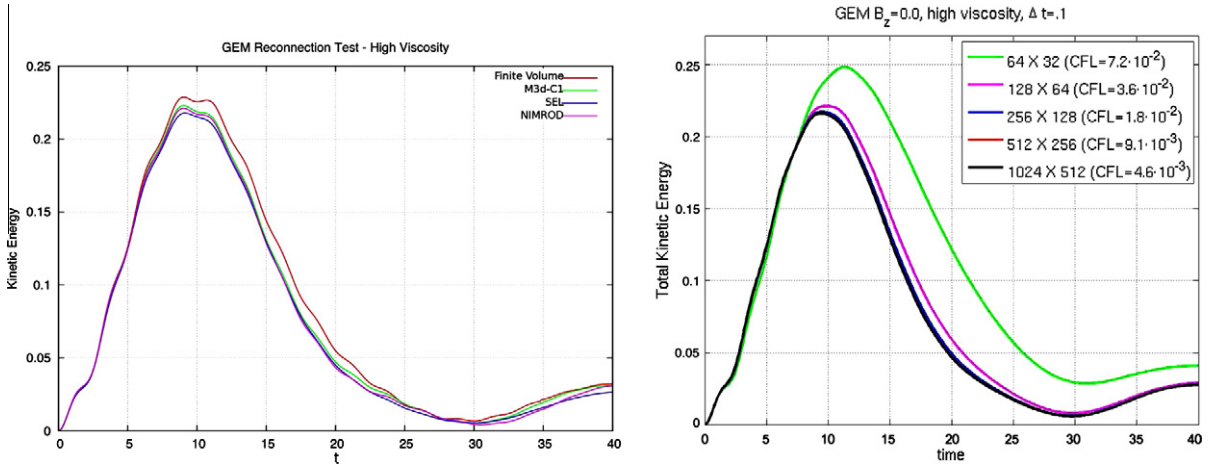

Fig. 4. Kinetic energy for high viscosity case with zero magnetic guide field with four other codes (left) and a spatial convergence study (right).

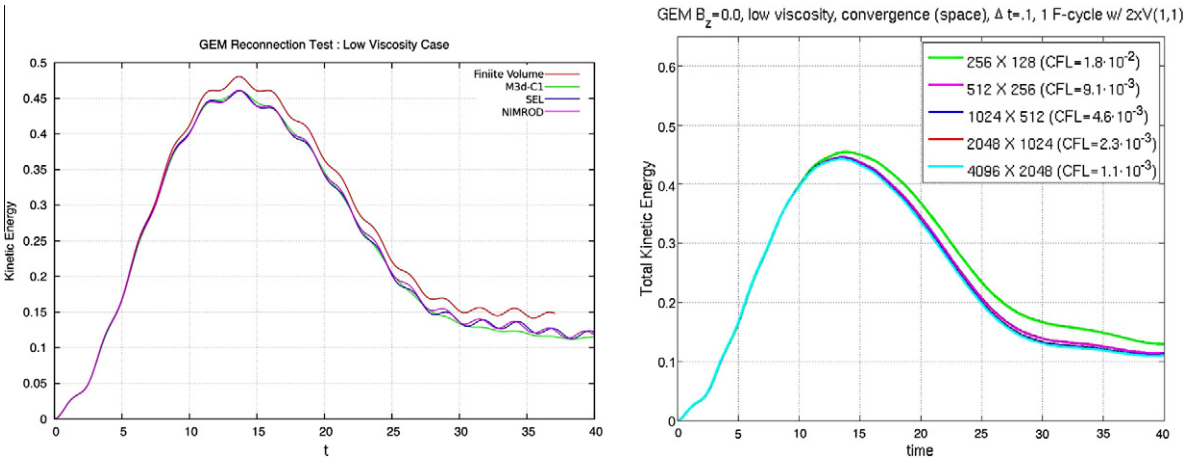

Fig. 5. Kinetic energy for low viscosity case with zero magnetic guide field with four other codes (left) and a spatial convergence study (right).

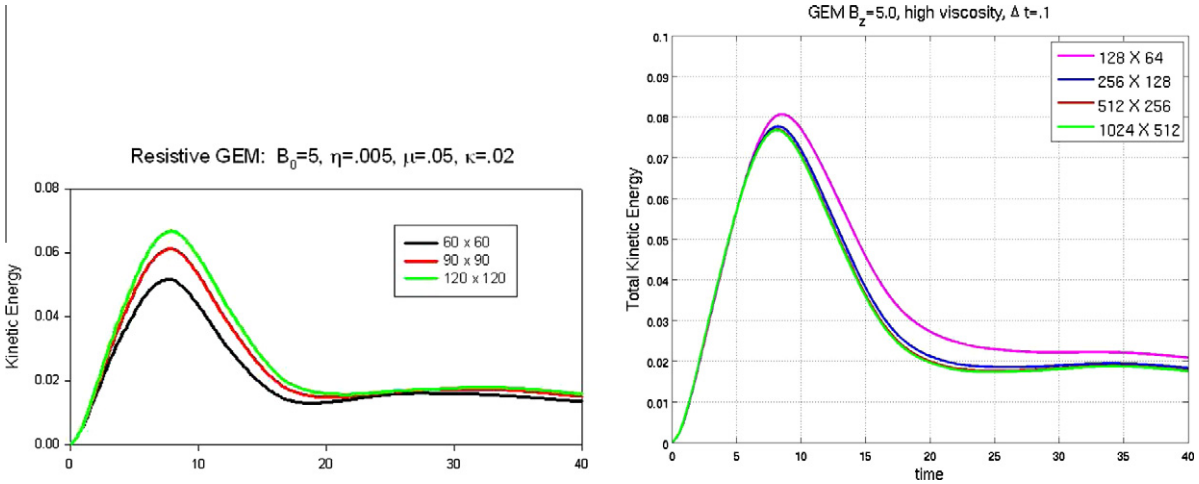

Fig. 6. Kinetic energy for high viscosity case with magnetic guide field with a finite element code (left) and a spatial convergence study (right).

Fig. 7 (left) shows the error as a function of spatial discretization (N, the number of cells in the long dimension) for several quantities of interest: (1) peak kinetic energy, (2) kinetic energy at $T = 40.0$, (3) reconnection flux rate at $T = 40.0$, and (4) reconnection rate at $T = 40.0$, for the low viscosity case with zero guide (B_z) field. Numerically, we do not preserve the solenoidal property of the magnetic field $\nabla \cdot B = 0$ exactly, even though the discretized system does. Errors in preserving the divergence property arise solely from the algebraic error. Numerical evidence in Fig. 7 (right) shows that these errors decrease with mesh refinement. This demonstrates second-order convergence on problems with up to a quarter of a billion equations ($8192 \times 4096 \times 8$). Observing second-order accuracy in space proves that our multigrid solver is reducing the algebraic error to the level of the discretization error, in that otherwise the algebraic error would start to dominate the solution and the error would level off or even increase as the problem size increased. Fig. 8 shows similar convergence studies for two other viscosity cases: (1) the high viscosity case with zero guide (B_z) field and (2) the low viscosity case with $B_z = 5$. These studies go up to grids with 4 billion and 1 billion equations, respectively, and Richardson extrapolation was used to construct approximations to the exact solutions, which were subsequently used to compute these errors. In the low viscosity, large guide field case we measure convergence at $T = 28.0$, which is about the time of the minimum kinetic energy.

Fig. 9 (right) shows the residual history, using the infinity norm, of ten F-cycle iterations on the high viscosity case. With the goal of solving the system with algebraic error on the order of the incremental truncation error, we would expect to see the rate of reduction of the algebraic error to increase by a factor of four in each successive halving of the mesh size, given that we use a second-order method. To the degree that the residual norm can be used as a proxy for the norm of the error this data does show a reduction rate of approximately three to four in the early iterations, as one would expect.

Fig. 4 (right), Fig. 5 (right), Fig. 6 (right), and Fig. 9 (left) show spatial convergence studies for the four cases in this study. From this data we conclude that a 512×256 grid is adequate for testing the ability of our solver to take large time steps.

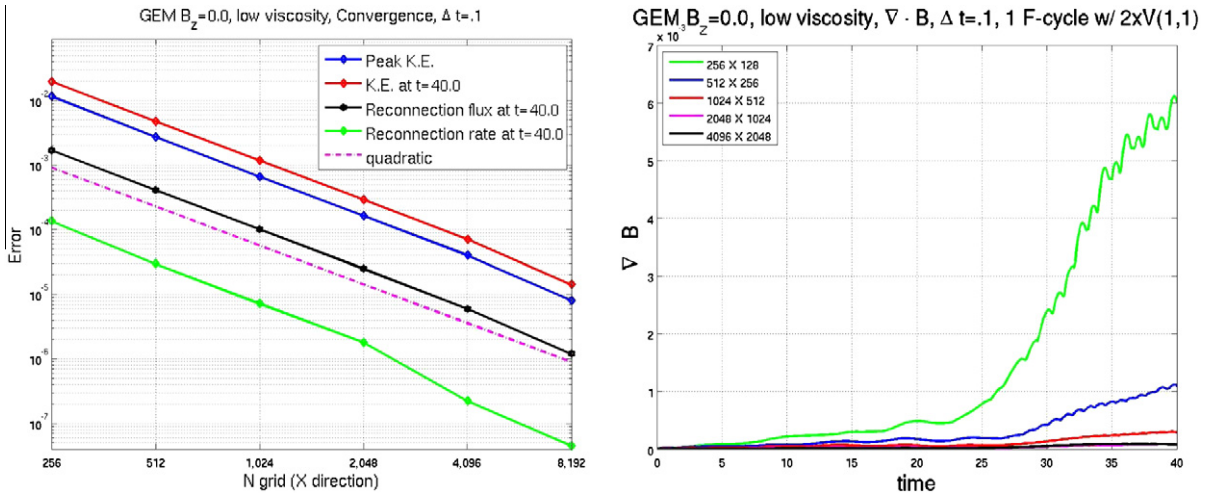

Fig. 7. Verification of second-order spatial accuracy. Error in peak kinetic energy and, kinetic energy, reconnection flux rate and reconnection rate at $T = 40.0$ (left), convergence of $\nabla \cdot B = 0$ (right).

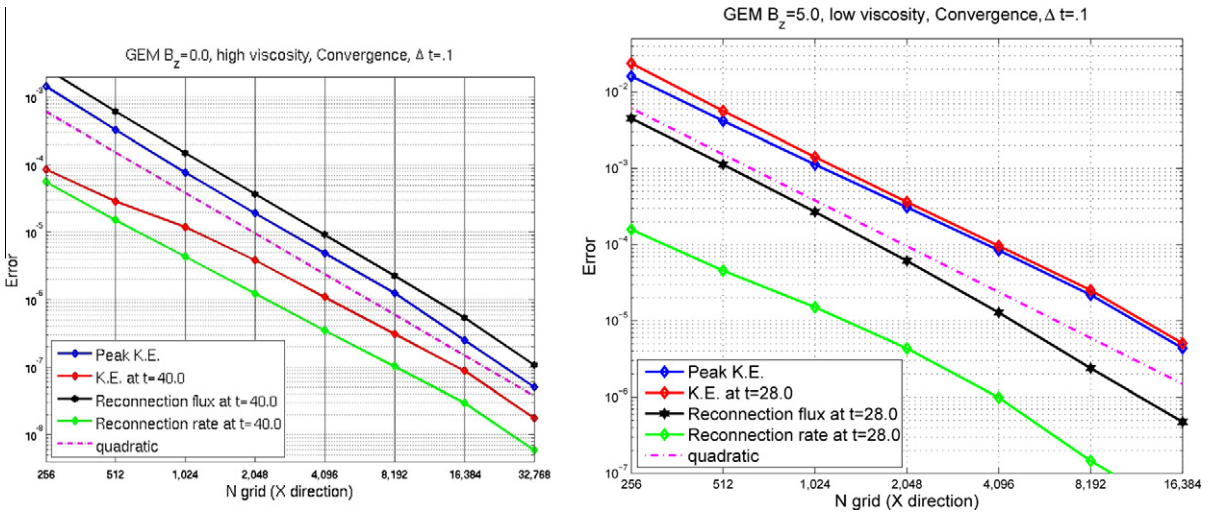

Fig. 8. Order of spatial accuracy. Error in peak kinetic energy and kinetic energy, reconnection flux rate and reconnection rate, high viscosity cases, $B_z = 0$ (left) and low viscosity $B_z = 5$ (right).

Fig. 10 shows time-step convergence studies using a 512×256 grid and $B_z = 0.0$. Note that reducing the time step not only decreases the truncation error, but also renders nonlinear systems that are more diagonally dominant, which results in increased accuracy from the iterative solver because the solver executes a fixed number of iterations instead of iterating until a desired reduction in the residual is achieved. Conversely, as the time step increases, the effectiveness of the solver decreases, leading to less-accurate solutions. The CFL limited time step for these problems is about $\Delta t = 0.009$ for the zero-guide field case and about $\Delta t = 0.002$ for the guide field case on the 512×256 grid. Fig. 10 (left) shows good accuracy at about $\Delta t = 0.25$, or about 25 times the CFL condition for the high viscosity case. This is about what is required to resolve the dynamics of the system, and the code is stable on times far larger than is useful for these dynamics. A common explicit method – second-order Runga–Kutta – has a work complexity of two per time step, and so we are nominally seeing a potential of about a 2.5x speed-up on this problem, give that our solver has a complexity of 18 work units or about 10x the explicit time-step cost. Fig. 11 shows time-step convergence studies using a 512×256 grid and $B_z = 5$. The high viscosity $B_z = 5$ case in Fig. 11 (left) shows very good time convergence at $\Delta t = 0.05$ with significant loss of accuracy at $\Delta t = 0.25$. The low viscosity case, Fig. 11 (right), similarly shows very good time convergence at $\Delta t = 0.05$ with significant loss of accuracy at $\Delta t = 0.25$. The chosen time step is about 50 times the CFL condition at $\Delta t = 0.1$ with a 512×256 grid. These results show that with zero-guide field and high viscosity we are able to take time steps well beyond what is needed to resolve the dynamics of the system accurately, which is about $\Delta t = 0.25$, but with low viscosity and/or a guide field we are limited to time steps of about $\Delta t = 0.1$.

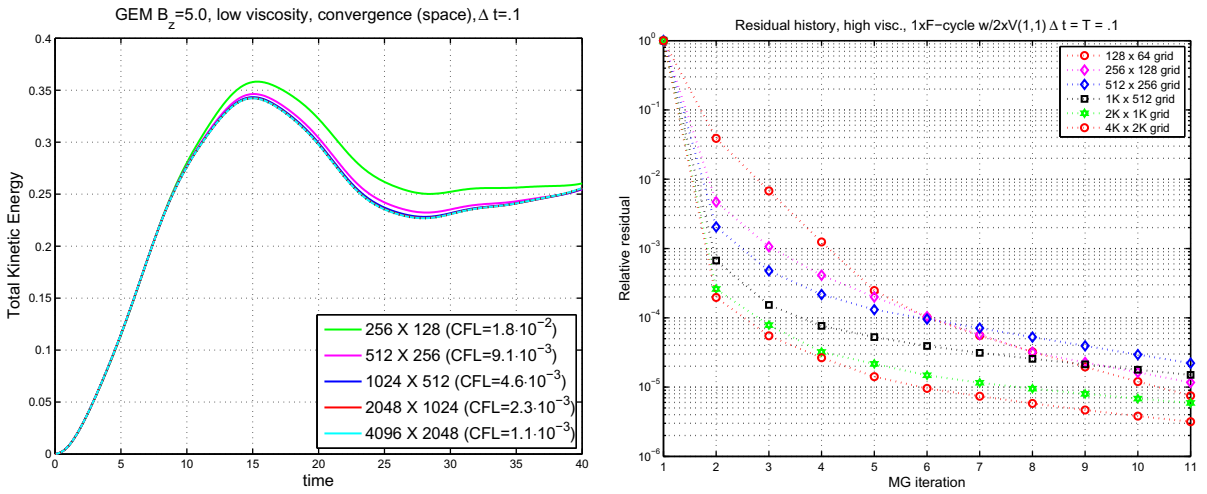

Fig. 9. Spatial convergence for kinetic energy for low viscosity case with magnetic guide field (left). Residual history of the first non-linear solve of the high viscosity, zero-guide field case.

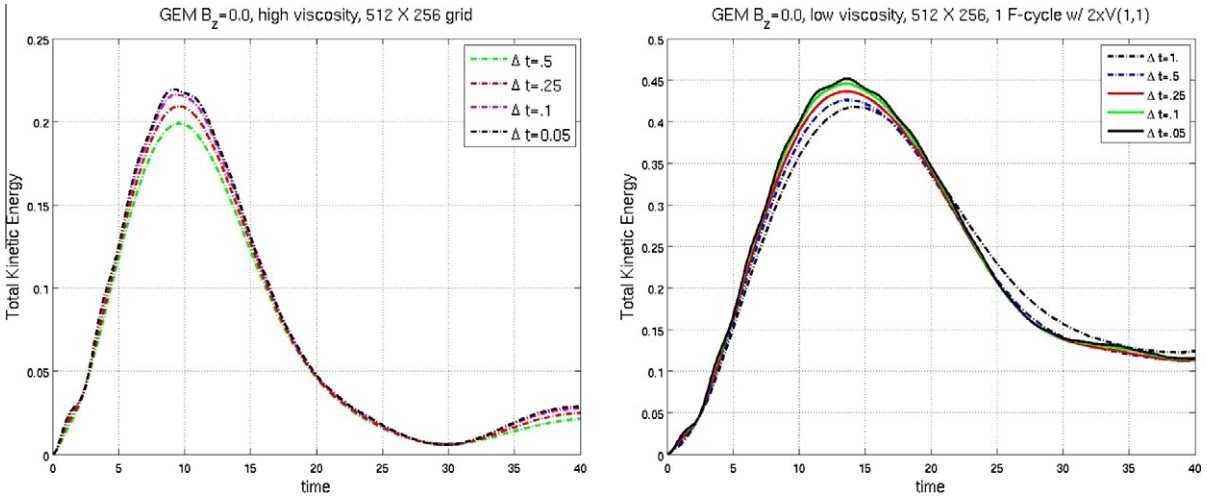

Fig. 10. Convergence in time for kinetic energy for zero-guide field: high viscosity (left) and low viscosity (right).

4.4. Additional numerical studies

This section considers the case of very low viscosity and resistivity with Re = 10,000 and S = 10,000, and a zero-guide field. Again, one F-cycle with two V (1,1) cycles at each level is used in each time step and $\Delta t = 0.1$. Fig. 12 (left) shows the kinetic energy as a function of time. Again we can see typical second-order convergence of the kinetic energy at $T = 20.0$. The norm of the error is an ideal measure of convergence though it is more difficult to compute. Fig. 12 (right) shows the infinity norm of the increment in the solution between successive grid refinements of the six active variables in this problem; i.e., a solution on a 256 × 128 grid is reduced to a 128 × 64 grid and the difference of this solution and the 128 × 64 grid solution is plotted, and so on. Given that we use a second-order method we expect the norm of these increments to fall quadratically with finer grids. The L_1 and L_2 norms also showed quadratic convergence, but are not presented here. These data confirm that we are indeed converging the solution asymptotically with second-order accuracy.

An alternative form of control of the solver is to monitor the residual and declare convergence at some tolerance value for the relative reduction in the residual (RTOL) at each time step. Fig. 13 shows a solver accuracy convergence study of the kinetic energy with several tolerance values (RTOL = 10^{-1}, 10^{-2}, and 10^{-4}), using V-cycles and F-cycles, as well as the single F-cycle with 2 V (1,1) cycles solver, on a 512 × 256 grid, $\Delta t = .1$, Re = 1,000 and S = 1,000, $B_z = 0$. This data shows that the kinetic energy is very accurate at earlier times and at time $T = 60.0$ is converged to within 1.9% with RTOL = 10^{-2} of the kinetic energy of the RTOL = 10^{-4} solve, and that the standard solver (on F-cycle) starts to loose accuracy at late times.

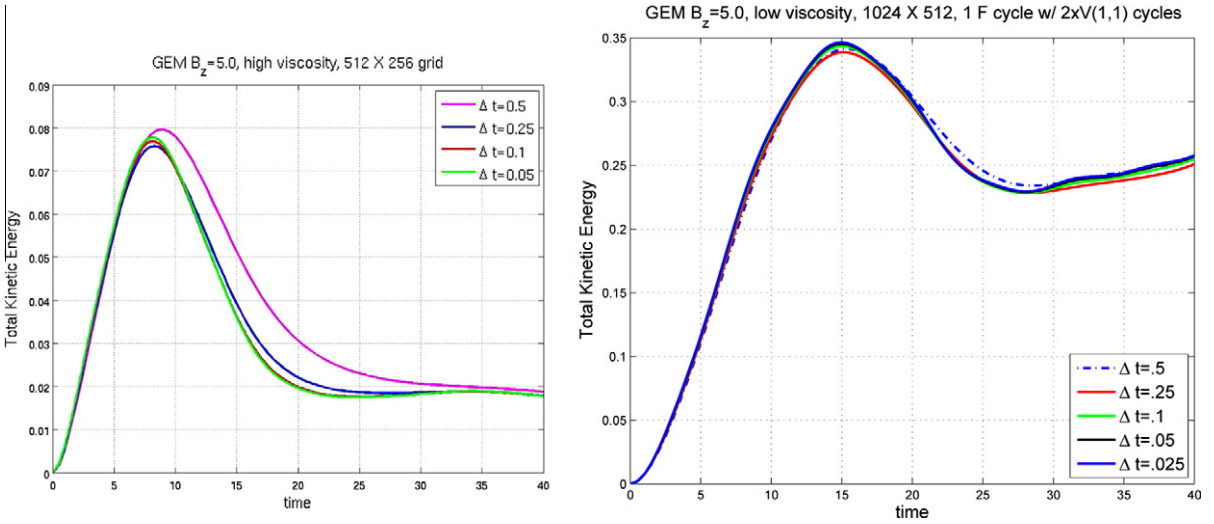

Fig. 11. Convergence in time for kinetic energy with a guide field of $B_z = 5$: high viscosity (left) and low viscosity (right).

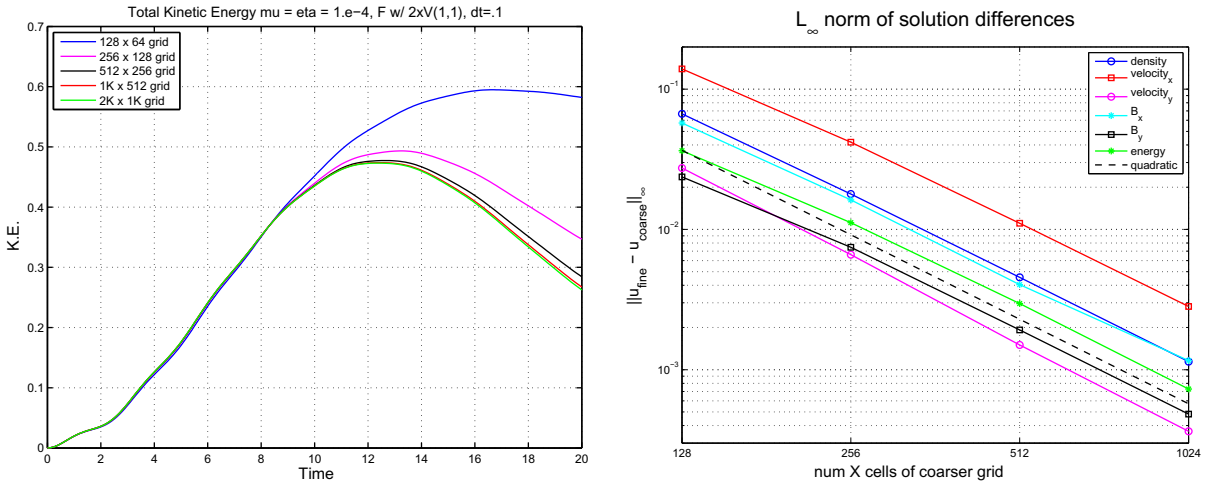

Fig. 12. Kinetic energy spacial convergence (left), Infinity norm of the solution increment (right), very low viscosity and resistivity case, $B_z = 0$.

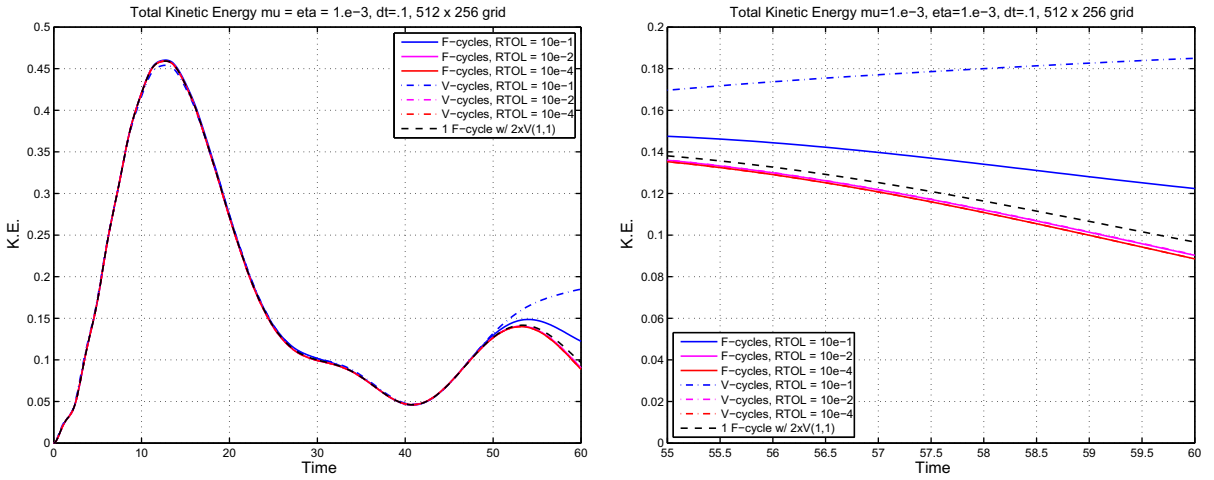

Fig. 13. Solver convergence of kinetic energy (left), detail (right), low viscosity and resistivity case, $B_z = 0$ on 512×256 grid.

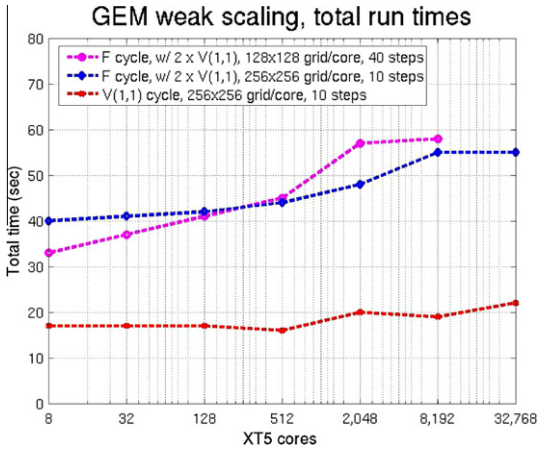

Fig. 14. Weak scaling of GEM reconnection problem on Cray XT-5.

4.5. Scaling results

We have shown that our methods are fast (solves with a few work units per time step) and scalable (asymptotically constant work per cell); in this section we show that these methods can also scale well in practice on a typical contemporary supercomputer – the "Jaguarpf" Cray XT-5 at ORNL. Fig. 14 shows the run times of our test problem as a function of the number of processor cores used. The number of cells is kept constant at either 128×128 grid per core or 256×256 grid per core. These are true weak-scaling in that the base case is run on only 8 cores, which is one node of the Cray XT-5, the minimal compute unit on this architecture. We see that the V-cycle is scaling almost perfectly and that the F-cycle is scaling well up to 32,768 cores, especially with the larger grid case.

5. Conclusion

We have developed a fully implicit nonlinear multigrid method for 2D resistive MHD and investigated the efficiency on the GEM reconnection problem with two different Reynolds numbers and with two magnetic guide fields. We have used a stable method for a low-order upwinding discretization as the smoother for a geometric multigrid solver for a second-order discretization via defect correction. We have demonstrated that a benchmark problem in magnetic reconnection can be solved with near textbook multigrid efficiency (e.g., about 18 work units per time step), even with the strong guild field that is needed for MHD simulation of magnetically confined fusion plasmas.

Our multigrid algorithm developed in this paper is matrix-free, i.e., there is no need to linearize and explicitly store a matrix, which is advantageous on advanced computer architectures where memory access is increasingly dominating the computational cost. The only data that need be stored are the primary data vectors of the algorithm (i.e., residuals, corrections, and solutions on all grids). Thus, this algorithm is well suited to utilize future computer architectures efficiently. Future work includes the development of stable high-order smoothers for high Reynolds number plasmas with techniques such as distributive Gauss–Seidel, multi-stage Jacobi, or KAPPA methods [25], and higher-order accurate time-integration solutions via defect correction.

Acknowledgments

This work was supported under the DOE SciDAC program (USDOE Contract No. DE-AC02-09CH11466) performed at Princeton Plasma Physics Laboratory, Princeton University, and Columbia University. This research used resources of the National Center for Computational Sciences at Oak Ridge National Laboratory, which is supported by the Office of Science of the US Department of Energy under Contract No. DE-AC05-00OR22725. This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231. We gratefully acknowledge verification data from the SciDAC Center for Extended MHD Modeling.

References

[1] I. Atlas, K. Burrage, A high accuracy defect correction multigrid method for the steady incompressible Navier–Stokes equations, J. Comput. Phys. 114 (1994) 227–233.

[2] J. Birn, J.F. Drake, M.A. Shay, B.N. Rogers, R.E. Denton, M. Hesse, M. Kuznetsova, Z.W. Ma, A. Bhattacharjee, A. Otto, P.L. Pritchett, Geospace environmental modeling (GEM) magnetic reconnection challenge, J. Geophys. Res. 106 (2001) 3715–3719.

[3] K. Böhmer, W. Gross, B. Schmitt, R. Schwarz, Defect corrections and Hartree–Fock method, in: K. Böhmer, H.J. Stetter (Eds.), Defect Correction Methods: Theory and Applications, Springer-Verlag, Vienna, 1984, pp. 193–209. Computing Suppl. 5.
[4] A. Brandt, Multi-level adaptive solutions to boundary value problems, Math. Comput. 31 (1977) 333–390.
[5] A. Brandt, Multi-level adaptive computations in fluid dynamics, AIAA J. 18 (1980) 1165–1172.
[6] A. Brandt, N. Dinar, Multigrid solutions to elliptic flow problems, in: S. Parter (Ed.), Numerical Methods for Partial Differential Equations, Academic Press, New York, 1979, pp. 53–147.
[7] L. Chacón, An optimal, parallel, fully implicit Newton–Krylov solver for three-dimensional viscoresistive magnetohydrodynamics, Phys. Plasmas 15 (5) (2008) 056103–056112.
[8] E. Dick, Second order formulation of a multigrid method for steady Euler equations through defect correction, J. Comput. Appl. Math. 35 (1991) 159–168.
[9] E. Dick, K. Riemslagh, Multi-staging of Jacobi relaxation to improve smoothing properties of multigrid methods for steady Euler equations, J. Comput. Appl. Math. 50 (1994) 241–254.
[10] M. Dryja, B.F. Smith, O.B. Widlund, Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions, SIAM J. Numer. Anal. 31 (6) (1994) 1662–1694.
[11] L.C. Dutto, W.G. Habashi, M. Fortin, An algebraic multilevel parallelizable preconditioner for large scale CFD problems, Comput. Meth. Appl. Mech. Eng. 149 (1997) 303–3018.
[12] N.M. Ferraro, S.C. Jardin, Calculations of two-fluid magnetohydrodynamic axisymmetric steady-states, J. Comput. Phys. 228 (20) (2009) 7742–7770.
[13] A.H. Glasser, X.Z. Tang, The SEL macroscopic modeling code, Comput. Phys. Commun. 164 (2004) 237–243.
[14] W. Hackbusch, Multi-grid Methods and Applications, Springer-Verlag, Berlin, 1985.
[15] P.W. Hemker, Defect correction and higher order schemes for the multigrid solution of the steady Euler equations, in: W. Hackbusch, U. Trottenberg (Eds.), Multigrid Methods II, Springer-Verlag, Berlin, 1986, pp. 149–165.
[16] A. Jameson, S. Yoon, Multigrid Solution of the Euler Equations Using Implicit Schemes, AIAA Paper, 85-0293, 1985.
[17] B. Koren, Low-diffusion rotated upwind schemes, multigrid and defect correction for steady, multi-dimensional Euler flows, in: W. Hackbusch, U. Trottenberg (Eds.), Multigrid Methods III, International Series of Numerical Mathematics, vol. 98, Birkhäuser, Basel, 1991, pp. 265–276.
[18] Wei Liao, Boris Diskin, Yan Peng, Li-Shi Luo, Textbook-efficiency multigrid solver for three-dimensional unsteady compressible Navier–Stokes equations, J. Comput. Phys. 227 (15) (2008) 7160–7177.
[19] D.J. Mavriplis, Directional Agglomeration Multigrid Techniques for High-Reynolds number Viscous Flows, Technical Report 98-7, Institute for Computer Applications in Science and Engineering Mail Stop 403, NASA Langley Research Center, Hampton, VA 23681-0001, 1998.
[20] D.R. Reynolds, R. Samtaney, C.S. Woodward, Operator-based preconditioning of stiff hyperbolic systems, SIAM J. Sci. Comput. 32 (1) (2010) 150–170.
[21] D.R. Reynolds, R. Samtaney, C.S. Woodward, A fully implicit numerical method for single-fluid resistive magnetohydrodynamics, J. Comp. Phys. 219 (2006) 144–162.
[22] R. Samtaney, P. Colella, T.J. Ligocki, D.F. Martin, S.C. Jardin, An adaptive mesh semi-implicit conservative unsplit method for resistive MHD, J. Phys. Conf. Ser. (2005) 40–48.
[23] C.R. Sovinec, A.H. Glasser, T.A. Gianakon, D.C. Barnes, R.A. Nebel, S.E. Kruger, S.J. Plimpton, A. Tarditi, M.S. Chu, The NIMROD Team, Nonlinear magnetohydrodynamics with high-order finite elements, J. Comp. Phys. 195 (2004) 355.
[24] H.R. Strauss, A. Pletzer, W. Park, S.C. Jardin, J. Breslau, L. Sugiyama, MHD simulations with resistive wall and magnetic separatrix, Comput. Phys. Commun. 164 (1-3) (2004) 40–45.
[25] U. Trottenberg, C.W. Oosterlee, A. Schüller, Multigrid, Academic Press, London, 2000.