A Constrained-Gradient Method to Control Divergence Errors in Numerical MHD

Philip F. Hopkins[1]∗

[1]*TAPIR & The Walter Burke Institute for Theoretical Physics, Mailcode 350-17, California Institute of Technology, Pasadena, CA 91125, USA*

ABSTRACT

In numerical magnetohydrodynamics (MHD), a major challenge is maintaining $\nabla \cdot \mathbf{B} = 0$. Constrained transport (CT) schemes achieve this but have been restricted to specific methods. For more general (meshless, moving-mesh, ALE) methods, "divergence-cleaning" schemes reduce the $\nabla \cdot \mathbf{B}$ errors; however they can still be significant and can lead to systematic errors which converge away slowly. We propose a new constrained gradient (CG) scheme which augments these with a projection step, and can be applied to any numerical scheme with a reconstruction. This iteratively approximates the least-squares minimizing, globally divergence-free reconstruction of the fluid. Unlike "locally divergence free" methods, this actually minimizes the numerically unstable $\nabla \cdot \mathbf{B}$ terms, without affecting the convergence order of the method. We implement this in the mesh-free code GIZMO and compare various test problems. Compared to cleaning schemes, our CG method reduces the maximum $\nabla \cdot \mathbf{B}$ errors by $\sim 1-3$ orders of magnitude ($\sim 2-5$ dex below typical errors if no $\nabla \cdot \mathbf{B}$ cleaning is used). By preventing large $\nabla \cdot \mathbf{B}$ at discontinuities, this eliminates systematic errors at jumps. Our CG results are comparable to CT methods; for practical purposes, the $\nabla \cdot \mathbf{B}$ errors are eliminated. The cost is modest, $\sim 30\%$ of the hydro algorithm, and the CG correction can be implemented in a range of numerical MHD methods. While for many problems, we find Dedner-type cleaning schemes are sufficient for good results, we identify a range of problems where using only Powell or "8-wave" cleaning can produce order-of-magnitude errors.

Key words: methods: numerical — hydrodynamics — instabilities — turbulence — cosmology: theory

1 INTRODUCTION

Magnetohydrodynamics (MHD) is essential to many physical problems, and (because the equations are non-linear) often requires numerical simulations. But this poses unique challenges. Naive discretizations of the MHD equations lead to violations of the "divergence constraint" ($\nabla \cdot \mathbf{B} = 0$); unfortunately, certain errors related to non-zero $\nabla \cdot \mathbf{B}$ are numerically unstable (they corrupt the solution even at infinite resolution). As such, many methods have been developed to control them.

The CT method of Evans & Hawley (1988), and related vector-potential/flux-central difference methods can conserve an initial $\nabla \cdot \mathbf{B} = 0$ to machine precision each timestep; however, while it may be possible to extend these methods in principle (see Mocz et al. 2014), it has thus far only been practical to implement for real problems in regular, fixed-grid (Eulerian) type schemes. It is also (often) computationally expensive.

For other (e.g. mesh-free, moving-mesh, arbitrary Lagrangian-Eulerian, smoothed-particle) methods, "divergence-cleaning" schemes are popular. The Powell et al. (1999) or "8-wave" cleaning simply subtracts the unstable $\nabla \cdot \mathbf{B}$ terms from the equation of motion (for a discussion of the instability, see Brackbill & Barnes 1980; Tóth 2000; Yang & Fengyan 2015); this cures the instability and restores Galilean invariance, but does not actually reduce $\nabla \cdot \mathbf{B}$ (it only "transports" the errors). Many studies have shown that certain types of problems, treated only with this method, will converge to the wrong solution (Tóth 2000; Mignone & Tzeferacos 2010; Mocz et al. 2014; Hopkins & Raives 2016); moreover the subtraction necessarily violates momentum conservation, so one would like to minimize the subtracted terms. More sophisticated cleaning schemes have been proposed; many of which follow Dedner et al. (2002) and add a scalar-field and set of source

terms which transport the divergence in waves and damp it (and correct behavior in shock jumps). However, this still requires a finite "response time" to damp $\nabla \cdot \mathbf{B}$ (so may act "too slowly" in discontinuities), and is reactive (dissipating, rather than preventing errors); as such it is less than ideal. Nevertheless, these schemes have been applied across a wide range of methods.

Alternatively, projection schemes following Brackbill & Barnes (1980) take the solution at each timestep, and project it onto a globally divergence-free basis (or equivalently, solve for the divergence-free component of the fluxes, and subtract off the other components).[1] This can reduce $\nabla \cdot \mathbf{B}$ below a desired tolerance at each timestep, acts "instantly," and as shown in Tóth (2000) preserves the convergence order of the method. However, it is expensive (requires a global sparse matrix inversion every timestep), can become unacceptably inefficient when adaptive/hierarchical timesteps and/or non-regular mesh geometries are used, is not compatible with arbitrary slope/flux limiters, and the inversion itself can become unstable under certain circumstances. Hence the application of these methods has been limited.[2]

In this paper, therefore, we propose a hybrid constrained-gradient (CG) scheme which combines some advantages of the

---

[1] This is generally done by representing $\mathbf{B}$ as $\mathbf{B} = \nabla \times \mathbf{A} + \nabla \phi$, solving for the vector/scalar fields $\mathbf{A}$ and ϕ, subject to some constraints (e.g. ϕ is constrained to minimize its least-squares value/integral over volume), then taking $\mathbf{B} \rightarrow \mathbf{B} - \nabla \phi$.

[2] Of course, many further divergence-control schemes have been proposed (see e.g. Swegle et al. 1995; Monaghan 2000; Børve et al. 2001; Maron & Howes 2003; Price & Rosswog 2006; Rosswog & Price 2007; Price & Bate 2008; Dolag & Stasyszyn 2009). However, many of these examples either fail to cure the numerical tensile instability, are zeroth-order inconsistent (so do not actually converge), are so diffusive and/or expensive as to be impractical, or cannot represent non-trivial magnetic field configurations. We will therefore not consider them further.

∗ E-mail:phopkins@caltech.edu

schemes above. We implement this in the multi-method Lagrangian MHD code GIZMO,[3] and test the scheme in a wide range of problems.

2 NUMERICAL METHODOLOGY

2.1 The Problem

In most finite-volume Godunov-type methods, the MHD equations are a set of hyperbolic partial differential conservation equations for an element (particle or cell) i, surrounded by elements j, which take the discrete form

$$\frac{d}{dt}(V\,\mathbf{U})_i + \sum_j \mathbf{F}_{ij} \cdot \mathbf{A}_{ij} = (V\,\mathbf{S})_i \qquad (1)$$

where V is the element volume, $\mathbf{U} = (\rho, \rho\mathbf{v}, \rho e, \mathbf{B}, \rho\psi)$ is a vector of primitive variables (mass/momentum/energy density, magnetic field, scalar fields), $\mathbf{S}$ is a vector of source terms, $\mathbf{A}_{ij}$ is an oriented "effective face area" defining the interaction surface between the elements, and $\mathbf{F}_{ij}$ is the relevant flux, computed by solving the appropriate Riemann problem at the face.[4]

The inputs to the Riemann problem at $\mathbf{A}_{ij}$ are the reconstructed quantities $\mathbf{U}_R$ and $\mathbf{U}_L$, which are the extrapolated values of $\mathbf{U}$ on the "i-side" and "j-side" of the face, respectively. In second-order methods, if the point at which the Riemann problem is solved (usually the midpoint of the face) lies at coordinates $\mathbf{x}_{ij}$, then $\mathbf{U}_R = \mathbf{U}'_i + \tilde{\phi}\,(\nabla\otimes\mathbf{U})_i \cdot (\mathbf{x}_{ij} - \mathbf{x}_i)$, where $(\nabla\otimes\mathbf{U})_i$ is the gradient tensor calculated at element i, $\tilde{\phi}$ is an appropriate slope-limiter which restricts the gradient values to prevent the creation of new maxima/minima, and $\mathbf{U}'_i$ denotes the value of $\mathbf{U}_i$ which is appropriately time-centered for the numerical time integration scheme.

In MHD, we wish to preserve $\nabla\cdot\mathbf{B} = 0$. Since there are an infinite number of valid definitions of the *discrete* gradient operator, it is impossible for any non-trivial field configuration to satisfy this under all operators. The *relevant* definition of $\nabla\cdot\mathbf{B}$, in most second-order finite-volume schemes, is something like

$$(V\,\nabla\cdot\mathbf{B})_i \equiv -\frac{1}{2}\sum_j \left(\mathbf{B}_R + \mathbf{B}_L + \frac{\psi_L - \psi_R}{c_{h,ij}}\,\hat{\mathbf{A}}_{ij} \right) \cdot \mathbf{A}_{ij} \qquad (2)$$

where the (optional) ψ terms arise from divergence-cleaning schemes such as that in Dedner et al. (2002), following a source function proportional to $(V\,\nabla\cdot\mathbf{B})_i$. The averaging of $\mathbf{B}_R$ and $\mathbf{B}_L$ appears because the one-dimensional Riemann problem requires that the $\mathbf{B}$ component normal to the face be constant. This definition represents the numerically unstable terms, which are subtracted in cleaning schemes (Powell et al. 1999; Dedner et al. 2002); also, since this represents a surface integral, maintenance of $(V\,\nabla\cdot\mathbf{B})_i = 0$ according to this definition is physically equivalent to magnetic flux conservation. Ideally, we would always have $\mathbf{B}_R = \mathbf{B}_L$ and $(V\,\nabla\cdot\mathbf{B})_i = 0$ (hence also $\psi = 0$). This is exactly what CT schemes try to ensure.

Note that variations of Eq. 2 are possible, and can (if carefully defined) represent valid gradient definitions. We will use Eq. 2 as the basis of our subsequent derivations because it is common to

many codes, and specifically it is the definition of relevance for the GIZMO code which we will use for our tests. However it is straightforward to modify our subsequent derivations for a modified form of Eq. 2.

2.2 Locally Divergence-Free Methods

We have some freedom in the choice of discrete approximation for the $\mathbf{B}$-field gradients, $\nabla\otimes\mathbf{B}$. For second-order methods, we must choose a definition such that the errors in the reconstruction scale $\propto h^2$ (where h is the linear element size) in smooth flows, but this still allows considerable flexibility.

To make progress, we will adopt the gradient estimator in GIZMO, a moving least-squares estimator. For a scalar f, this is

$$(\nabla f)_i^\alpha = \sum_j (f_j - f_i)\left(\mathbf{W}_i^{-1}\right)^{\alpha\beta} (\mathbf{x}_j - \mathbf{x}_i)^\beta \,\omega_j(\mathbf{x}_i) \qquad (3)$$

$$\mathbf{W}_i^{\alpha\beta} \equiv \sum_j (\mathbf{x}_j - \mathbf{x}_i)^\alpha (\mathbf{x}_j - \mathbf{x}_i)^\beta \,\omega_j(\mathbf{x}_i) \qquad (4)$$

here we assume an Einstein summation convention over the indices β corresponding to the spatial dimensions, and $\omega_j(\mathbf{x}_i)$ is an (arbitrary) weight function defined in Paper I. This estimator is second-order accurate for an arbitrary mesh configuration, minimizes the (weighted) least-squares deviation $\sum_j \omega_j |f_i + \nabla f_i \cdot (\mathbf{x}_j - \mathbf{x}_i) - f_j|^2$, and has been applied in a wide range of different numerical methods (see e.g. Oñate et al. 1996; Kuhnert 2003; Maron & Howes 2003; Luo et al. 2008; Lanson & Vila 2008).

It is straightforward to constrain this to obtain the locally divergence-free solution: i.e. $(\nabla\otimes\mathbf{B})_i$ such that $(\nabla\cdot\mathbf{B})_i \equiv \sum_k (\nabla\otimes\mathbf{B})_{i,kk} = 0$. In the least-squares formulation, this is just a constrained least-squares problem – we seek the matrix $(\nabla\otimes\mathbf{B})_i$ constrained to have $(\nabla\cdot\mathbf{B})_i = 0$ which minimizes the (weighted) square deviation of $\mathbf{B}_i + (\mathbf{x}_j - \mathbf{x}_i)\cdot(\nabla\otimes\mathbf{B})_i - \mathbf{B}_j$.

A similar approach is to calculate the gradient projected onto a set of divergence-free basis functions. The matrix formulation above implicitly adopts the Cartesian polynomial basis functions $(1, x, y, z, x^2, xy, xz, ..)$, but a divergence-free basis can be chosen instead and used for the reconstruction.

The problem with both of these is as follows. If we temporarily ignore the slope-limiter and ψ terms in Eq. 2, then $(V\,\nabla\cdot\mathbf{B})_i$ is just

$$(V\,\nabla\cdot\mathbf{B})_i = -\frac{1}{2}\sum_j \Big[\mathbf{B}_i + (\nabla\otimes\mathbf{B})_i \cdot (\mathbf{x}_{ij} - \mathbf{x}_i) \qquad (5)$$
$$+ \mathbf{B}_j + (\nabla\otimes\mathbf{B})_j \cdot (\mathbf{x}_{ij} - \mathbf{x}_j)\Big] \cdot \mathbf{A}_{ij} + \mathcal{O}\left(\tilde{\phi}, \psi\right)$$

It is obvious from this expression that ensuring $(\nabla\cdot\mathbf{B})_i = 0$ does not ensure $(V\,\nabla\cdot\mathbf{B})_i = 0$; in fact, it does not even necessarily decrease $(V\,\nabla\cdot\mathbf{B})_i$.[5]

2.3 An Approximate, Globally Divergence-Free Method

Still, these locally divergence-free methods suggest a solution. Assume, for now, that we are estimating the gradient $(\nabla\cdot\mathbf{B})_i$ of element i, and all other gradients in the system have been determined;

---

[3] A public version of this code is available at `http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html`. Users are encouraged to modify and extend the capabilities of this code; the development version of the code is available upon request from the author.

[4] Throughout, "$\otimes$" denotes the outer product, "$\cdot$" the inner (dot) product, and "$:$" the double-dot product $\mathbf{A}:\mathbf{B} = \sum_{ij} A_{ij} B_{ij}$. We use $\sum_{ij}$ to denote double-summation $\sum_{ij} \equiv \sum_i \sum_j$.

[5] It is easy to verify this. Consider a trivial case: a 2D, perfectly regular lattice of equally-spaced elements so $|\mathbf{A}_{ij}| = A$ and $|\mathbf{x}_{ij} - \mathbf{x}_i| = \delta x$, and take $\mathbf{x}_i = 0$. Assuming $(\nabla\cdot\mathbf{B})_i = 0$, Eq. 5 simplifies to $(V\,\nabla\cdot\mathbf{B})_i = -A[(B_x(\delta x, 0) - B_x(-\delta x, 0) + B_y(0, \delta x) - B_y(0, -\delta x)]/2$. Assume $B_y = f(y)$, so $B_x = B_{x,0}(y) - (df/dy)x$; then $(V\,\nabla\cdot\mathbf{B})_i = A[(df/dy)\delta x + (f(\delta x) - f(-\delta x))/2] \neq 0$ for any non-linear f. Or simply assume a noisy, but constant-mean field, such that $(\nabla\otimes\mathbf{B})_i = 0$; then $(\nabla\cdot\mathbf{B})_i = 0$, but $(V\,\nabla\cdot\mathbf{B})_i \propto \sum_j (\mathbf{B}_i + \mathbf{B}_j)\cdot\mathbf{A}_{ij}$, which can only vanish for special configurations of $\mathbf{B}_j$ and $\mathbf{A}_{ij}$.

also for now neglect the slope-limiter and ψ terms so $(V \nabla \cdot \mathbf{B})_i$ follows Eq. 5. Then we see that we can, in fact, ensure $(V \nabla \cdot \mathbf{B})_i = 0$, provided that $(\nabla \otimes \mathbf{B})_i$ satisfies:

$$(\nabla \otimes \mathbf{B})_i : \sum_j (\mathbf{x}_{ij} - \mathbf{x}_i) \otimes \mathbf{A}_{ij} = \tag{6}$$

$$- \sum_j \left[\mathbf{B}_i + \mathbf{B}_j + (\nabla \otimes \mathbf{B})_j \cdot (\mathbf{x}_{ij} - \mathbf{x}_j) \right] \cdot \mathbf{A}_{ij}$$

Or, in component form,

$$\sum_{ab} (\nabla \otimes \mathbf{B})_i^{ab} Q^{ab} = \sum_{ab} (\partial B^a / \partial x^b)_i Q^{ab} = S_0 \tag{7}$$

$$\mathbf{Q} \equiv \sum_j (\mathbf{x}_{ij} - \mathbf{x}_i) \otimes \mathbf{A}_{ij} \tag{8}$$

$$S_0 \equiv - \sum_j \left[\mathbf{B}_i + \mathbf{B}_j + (\nabla \otimes \mathbf{B})_j \cdot (\mathbf{x}_{ij} - \mathbf{x}_j) \right] \cdot \mathbf{A}_{ij} \tag{9}$$

This is a single scalar equation: so it only constrains one degree of freedom of $(\nabla \otimes \mathbf{B})_i$.

Say that our "preferred" gradient, in the absence of this constraint, is $(\nabla \otimes \mathbf{B})_{i,0}$. This is the gradient (accurate to the desired level of reconstruction) calculated by whatever default method, before any consideration of the constraint. Then define

$$(\nabla \otimes \mathbf{B})_i = (\nabla \otimes \mathbf{B})_{i,0} + \mathbf{G} \tag{10}$$

where $\mathbf{G}$ is a correction term that ensures $(\nabla \otimes \mathbf{B})_i$ satisfies Eq. 6. We would like to make $(\nabla \otimes \mathbf{B})_i$ as close to $(\nabla \otimes \mathbf{B})_{i,0}$ as possible, so we choose the tensor $\mathbf{G}$ which satisfies Eq. 6 while minimizing some "penalty function," which in this paper we take to be the least-squares deviation

$$f_{\mathrm{penalty}} \equiv \left| (\nabla \otimes \mathbf{B})_i - (\nabla \otimes \mathbf{B})_{i,0} \right|^2 = |\mathbf{G}|^2 = \sum_{ab} \left(G^{ab} \right)^2 \tag{11}$$

This gives the solution

$$\mathbf{G} = -\mathbf{Q} \left(\frac{S_0 + (\nabla \otimes \mathbf{B})_{i,0} : \mathbf{Q}}{\mathbf{Q} : \mathbf{Q}} \right) \tag{12}$$

$$G^{ab} = -Q^{ab} \left(\frac{S_0 + \sum_{cd} (\nabla \otimes \mathbf{B})_{i,0}^{cd} Q^{cd}}{\sum_{cd} (Q^{cd})^2} \right) \tag{13}$$

We note that while this changes the gradient from our "preferred" estimator, and will, of course, change the errors in the solution, the term $\mathbf{G}$ is explicitly minimized, and is sourced by non-zero $\nabla \cdot \mathbf{B}$ errors. In any well-designed numerical scheme, these errors are of the convergence order of the code – therefore, although this alters the gradients, it is easy to show that it does *not* change the convergence order.

There are two problems that prevent this from exactly providing $(V \nabla \cdot \mathbf{B})_i = 0$. (**1**) First, $\mathbf{G}$, and hence each $(\nabla \otimes \mathbf{B})_i$, depends on the neighboring gradients $(\nabla \otimes \mathbf{B}_j)$, through Eq. 9. (**2**) Second, slope-limiters and non-zero ψ will add terms to $\nabla \cdot \mathbf{B}$ which are not accounted for.

Issue (**1**) could be eliminated by solving *simultaneously* for all $(\nabla \otimes \mathbf{B})_i$. This is technically easy, but expensive (it amounts to a global sparse-matrix inversion every timestep), and "overkill," because issue (**2**) would still prevent $(V \nabla \cdot \mathbf{B})_i = 0$. A much simpler, and more computationally efficient solution is to iteratively calculate $(\nabla \otimes \mathbf{B})_i$:

$$(\nabla \otimes \mathbf{B})_i^{(n)} = (\nabla \otimes \mathbf{B})_{i,0} - \mathbf{Q} \left(\frac{S_0^{(n)} + (\nabla \otimes \mathbf{B})_{i,0} : \mathbf{Q}}{\mathbf{Q} : \mathbf{Q}} \right) \tag{14}$$

$$S_0^{(n)} \equiv - \sum_j \left[\mathbf{B}_i + \mathbf{B}_j + (\nabla \otimes \mathbf{B})_j^{(n-1)} \cdot (\mathbf{x}_{ij} - \mathbf{x}_j) \right] \cdot \mathbf{A}_{ij} \tag{15}$$
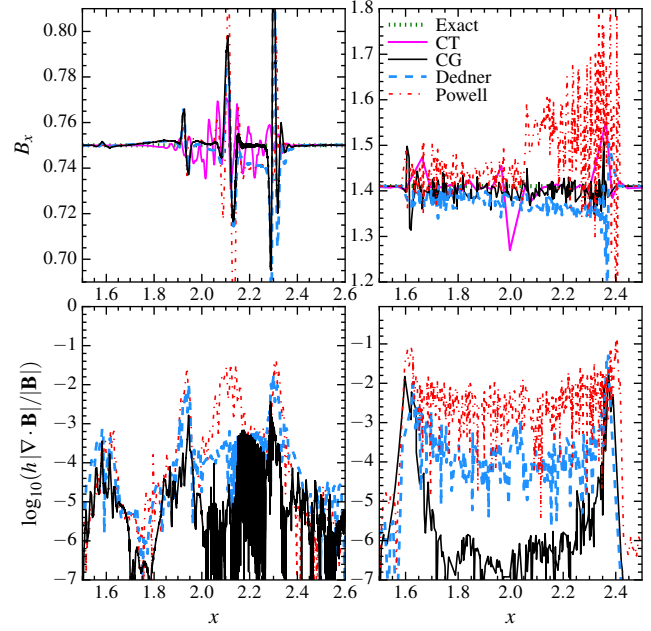
Figure 1. Brio-Wu (*left*) and Toth (*right*) shocktubes (§ 3.1), at times $t = 0.2$ and $t = 0.08$, respectively. The setups for these and all other test problems follow Hopkins & Raives (2016). The tubes are 2D, with ~ 256 elements across the $\hat{x}$ (defined as the direction of shock propagation), with the initial grid mis-aligned from $\hat{x}$. We show the $\hat{x}$ component of $\mathbf{B}$ (*top*), which should be constant, and $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$ (*bottom*), which measures the fractional magnitude of the magnetic divergence errors. All other fluid quantities (P, u, ρ, $\mathbf{v}$) tend to agree more closely between methods and are less sensitive to the divergence-control method. With no divergence control, catastrophic errors overwhelm any solution (see Fig. 2). Using only the Powell et al. (1999) "8-wave" cleaning ("Powell"), $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$ reaches ~ 0.1, large noise/oscillations appear, and the incorrect shock jump is recovered, producing a systematic offset in B_x (at $x \sim 2.1 - 2.3$ and $x \sim 2.0 - 2.4$, respectively). This offset does not decrease with resolution. Using the more sophisticated Dedner et al. (2002) cleaning reduces $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$ by $\sim 1 - 2$ dex, suppresses the oscillations, and dramatically reduces the systematic offset at the shock jump. However an offset still exists in both problems at the $\sim 2 - 5\%$-level, which converges away slowly ($\propto N_{1D}^{-0.5}$). Our new constrained-gradient (CG) method maintains $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i < 0.01$ at the discontinuities, and $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i \ll 10^{-4}$ elsewhere; it produces still-smaller oscillations, and most important, completely eliminates the systematic offset at the jump. Constrained transport (CT) maintains $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i \lesssim 10^{-14}$ here; however, because the shock is not exactly grid-aligned, oscillations in B_x still appear, comparable to CG.

where in the first-pass ($n = 1$), we take $(\nabla \otimes \mathbf{B})_j^{(0)}$ to be the value of $(\nabla \otimes \mathbf{B})_j$ from the most recent time/drift step, calculate the new $(\nabla \otimes \mathbf{B})_i^{(1)}$ for all elements, then use these values to calculate the updated $(\nabla \otimes \mathbf{B})_i^{(2)}$, etc.

In practice, we find that given the errors sourced by the slope-limiters, we converge to nearly best-case accuracy in just two iterations ($n = 2$). And since all the quantities here can be calculated in the same pass that is used to calculate $(\nabla \otimes \mathbf{B})_{i,0}$, the entire iteration series only requires one additional element sweep, compared to the "standard" method.

2.4 Complications: Dealing with Slope Limiters and Cleaning Terms

Issue (**2**) is more challenging. The ψ terms in Eq. 2 should be minimized along with $\nabla \cdot \mathbf{B}$, so they are not particularly problematic. We can include them explicitly in our constraint solution for $\mathbf{G}$, using the same iterative approach to account for the fact that ψ_L, ψ_R,
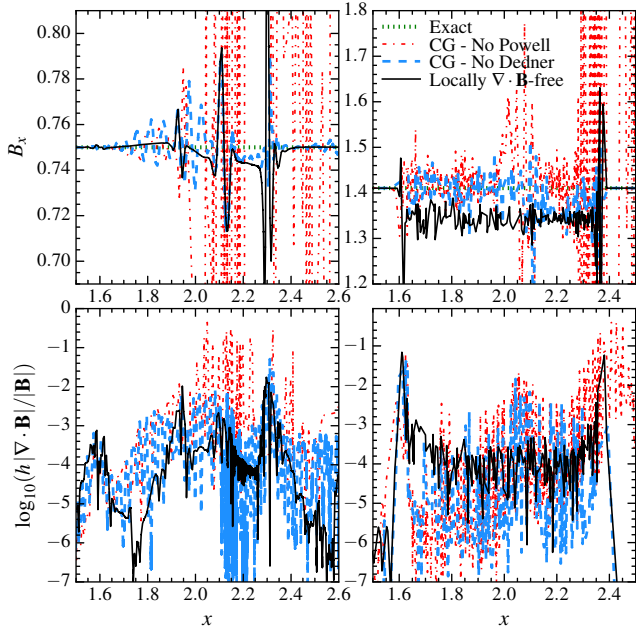
Figure 2. Shocktubes from Fig. 1, with alternative divergence-control schemes. If, instead of approximating the global, divergence-free reconstruction according to Eq. 6, we simply constrain the system to be "locally divergence free" (i.e. $\nabla \cdot \mathbf{B} = 0$ for the particle-centered gradient estimate; as § 2.2), we see there is essentially no reduction in the numerically problematic $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$ term compared to the Dedner et al. (2002) without this constraint, and the systematic shock jump errors are actually increased. If we apply our CG scheme without the Dedner et al. (2002) cleaning, we recover the mean solutions but see large oscillations since the terms driving corrections to the gradients are not being damped. If we apply our CG scheme but ignore the Powell et al. (1999) correction terms, the tensile instability appears and the oscillations grow to unacceptable levels.

and $c_{h,ij}$ themselves depend on the gradients in the problem. However, this "mixing" of ψ and $\mathbf{B}$ essentially defeats the purpose of the damping ψ terms, and can introduce more serious numerical instabilities in the (rare) cases where $\psi / c_h \gg |\mathbf{B}|$. We therefore leave them, since their purpose is to damp $\nabla \cdot \mathbf{B}$ where present. But we do reduce the contribution of the $\psi_L - \psi_R$ term in Eq. 2 by minimizing the least-squares deviation between ψ extrapolated from the i and j "sides" *at the face locations*, rather than at the particle-j locations (i.e. our preferred ψ gradient minimizes the squared deviation of $[\psi_i + (\nabla \psi)_i \cdot (\mathbf{x}_{ij} - \mathbf{x}_i)] - [\psi_j + (\nabla \psi)_j \cdot (\mathbf{x}_{ij} - \mathbf{x}_j)]$, rather than $\psi_i + (\nabla \psi)_i \cdot (\mathbf{x}_j - \mathbf{x}_i) - \psi_j$). This is numerically consistent at the same order and trivial to implement using the same matrix based least-squares formulation, and we find it slightly reduces the divergences and improves the cleaning accuracy, so we use it throughout.

To deal with the slope-limiters $\tilde{\phi}$, we take advantage of our iterative approach. Generally speaking, there are two types of slope-limiters in most of the methods of interest.

First, slope-limiters that are applied to the gradient after its calculation loop (and apply to all subsequent operations): $\tilde{\phi}$ is $\mathrm{MIN}(1, \tilde{\phi}')$ where $\tilde{\phi}'$ is chosen such that the reconstruction value of a field does not exceed the maximum/minimum neighbor values by more than some tolerance (Balsara 2004).[6] These are straight-

[6] In GIZMO, this takes the form:

$$\tilde{\phi}'_i \equiv \mathrm{MIN}\left[1, \beta_i \mathrm{MIN}\left(\frac{U_{ij,\mathrm{ngb}}^{\max} - U_i}{U_{ij,\mathrm{mid}}^{\max} - U_i}, \frac{U_i - U_{ij,\mathrm{ngb}}^{\min}}{U_i - U_{ij,\mathrm{mid}}^{\min}}\right)\right] \quad (16)$$
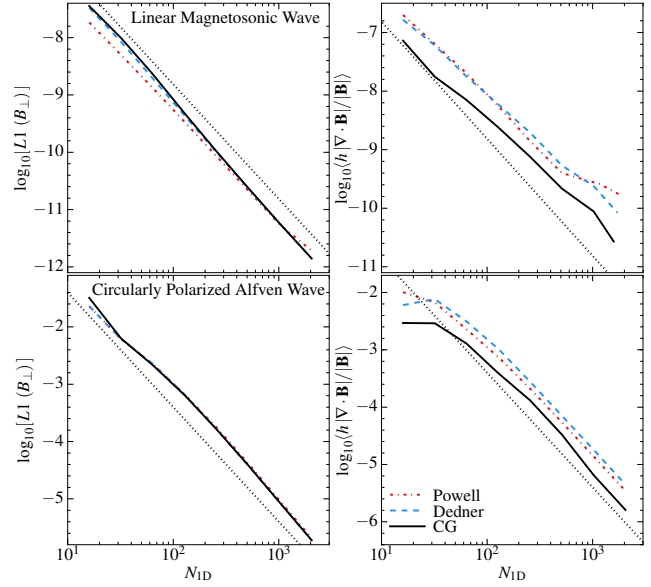

Figure 3. Convergence study in MHD wave tests (§ 3.2). *Top:* Linear magnetosonic wave: a linear 1D fast magnetosonic wave is propagated one wavelength; we then plot the $L1$ norm (*left*; mean absolute error relative to the analytic solution) in magnetic field ($B_\perp$, direction perpendicular to propagation), and mean absolute divergence error $\langle h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i \rangle$ (*right*; this is dominated by the largest errors in the domain, so the scaling is similar for $\mathrm{MAX}(h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i)$). *Bottom:* Circularly polarized Alfven wave test: here the wave is an exact *non-linear* solution of the MHD equations. The wave is evolved in 3D and is tilted by $\sim 27°$ relative to the *x*-axis, and errors are measured after is propagates five wavelengths. All errors are plotted as a function of the number of elements across the domain N_{1D}; dotted lines show second-order convergence ($L1 \propto N^{-2}$). We compare our CG, Dedner, and Powell methods; note that the CT results are from a different code with different convergence properties so a comparison here is not appropriate. In all cases, the methods here show roughly second-order convergence (slightly faster/slower for $B_\perp$ and $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$ in the linear wave test, but this is in the limit where the errors in some quantities approach floating-point accuracy). As expected, the CG method systematically reduces $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$ (even moreso if we consider the median $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$). At low resolution, this comes at a small cost in accuracy (slightly larger errors in $\mathbf{B}$), owing to the constrained reconstruction of the magnetic field. However this appears to converge away quickly.

forward: we calculate our "preferred" gradient and then apply this limiter, and treat this as the new "preferred" gradient. After correction, the new gradient may violate this condition, so we can (optionally) re-limit it, and treat this as the new "preferred" gradient, and iterate until convergence (this iteration is outside the neighbor loop so has negligible cost). This converges to the gradient satisfying the desired slope limiter which comes as close as possible to the desired CG-corrected gradient.

Second, another class of slope-limiters can (optionally) be additionally applied in pair-wise fashion between every interacting element pair in the flux computations; this ensures no local maxima/minima are created. Here the limiter $\tilde{\phi}_{ij}$ is unique to the ele-

where $U_{ij,\mathrm{ngb}}^{\max}$ and $U_{ij,\mathrm{ngb}}^{\min}$ are the maximum and minimum values of U_j among all neighbors j of the particle i, and $U_{ij,\mathrm{mid}}^{\max}$, $U_{ij,\mathrm{mid}}^{\min}$ are the maximum and minimum values (over all pairs ij of the j neighbors of i) of U re-constructed on the "i side" of the interface between particles i and j. The constant $\beta = 1 - 2$ depending on local particle order. In our iterative CG implementation, we first limit with the "normal" $\beta = \beta_0$, then correct the gradient, then re-limit only if it exceeds the slightly weaker limiter with $\beta = 2\beta_0$.

ment pair. We account for these limiters explicitly in our calculation of S_0: using both the current values of $(\nabla \otimes \mathbf{B})_i$ and $(\nabla \otimes \mathbf{B})_j$, we apply the limiters between each pair, and thus obtain a more accurate guess for the correction to $(\nabla \otimes \mathbf{B})_i$.

These approaches allow us to handle arbitrary slope limiters and still return some valid result. But it is easy to see that application of any slope limiter can, under some circumstances, disallow the corrected value of $(\nabla \otimes \mathbf{B})_i$ needed to actually ensure $(V \nabla \cdot \mathbf{B})_i = 0$. This is why our procedure ceases to significantly improve after a couple iterations. And clearly, a stricter slope-limiter prevents $(V \nabla \cdot \mathbf{B})_i$ correction under a wider range of circumstances. Therefore, when we implement our CG method, we "weaken" our normal pair-wise slope limiter for $\mathbf{B}$, to allow more flexible CG correction. This is important because, as we will show, if we do this and do not include any divergence-damping terms, it (unsurprisingly) produces large oscillations. Note, though, that we still retain the standard slope-limiter applied after the gradient calculation ($\tilde{\phi}'$), so the second slope limiter is somewhat redundant anyways, and we only alter the slope limiters for $\mathbf{B}$.

2.5 Implementation

We implement the CG method above in the code GIZMO (Hopkins 2015). GIZMO is a mesh-free, finite-volume Godunov code, built on the gravity solver and domain decomposition algorithms of GADGET-3 (Springel 2005). In Hopkins (2015); Hopkins & Raives (2016) we consider extensive surveys of test problems in both hydrodynamics and MHD (using the Dedner scheme) with this code, and demonstrate accuracy and convergence in good agreement with well-studied regular-mesh finite-volume Godunov methods. Because GIZMO is a multi-method code, it allows us to compare directly the effects of different divergence-control methods with an otherwise entirely identical code. Since we are not comparing hydro solvers here but the divergence control method, in what follows, we run GIZMO always in its MFM (Meshless Finite-Mass) mode, but we note that we have run several problems in the MFV (Meshless Finite-Volume) mode, and find nearly identical results (as expected from the comparisons in the methods paper). We have also implemented a limited, 2D version of CG in the public moving-mesh code FVMHD3D (Gaburov et al. 2012);[7] as expected from our previous comparisons, this is very similar to the GIZMO MFV results. For reasons shown below, when we implement our CG method, we still retain the Powell and Dedner source terms in the MHD equations, to deal with imperfect minimization of $(V \nabla \cdot \mathbf{B})_i$.

We have made public a version of the GIZMO code with the CG implementation used in this paper;[8] users interested in the details of our implementation (for example, the exact numerical form of the slope-limiters used, kernel weights used for the least-squares calculation, etc.) are encouraged to examine the source code.

3 TEST PROBLEMS

We now consider a series of test problems. Each of these has been studied in detail in Hopkins & Raives (2016), where we undertook a systematic comparison of different algorithms (MFM, MFV, SPH-MHD, and moving meshes, using the Dedner et al. 2002

divergence-cleaning scheme, and grid/AMR schemes using constrained transport). Therefore we will not describe them in detail here.

3.1 Shocktubes

Fig. 1 shows two shocktube standard shocktube tests: the subsonic, magnetically dominated Brio & Wu (1988) and super-sonic Tóth (2000) shocktubes. We compare different divergence-control schemes, with resolution $\approx 256 \times 56$ across the domain plotted (a 2D grid, with the shock propagating at an angle $\pi/6$ to the grid).[9]

As discussed at length in Hopkins & Raives (2016), with no divergence-control at all, the schemes are unstable and crash (developing negative pressures). The minimal correction to restore stability is the Powell et al. (1999) or "8-wave" cleaning; however this only subtracts the tensile terms from non-zero $\nabla \cdot \mathbf{B}$, it does not actually control $\nabla \cdot \mathbf{B}$; as a result we show therein that it produces incorrect shock jumps in B_x and u, which lead to the shock being in the wrong position over time. Most importantly, these are zeroth-order errors which *do not converge* away at any resolution. This is known from previous studies as well (Tóth 2000; Mignone & Tzeferacos 2010; Mocz et al. 2014).

Adding the Dedner et al. (2002) divergence-damping greatly reduces $\nabla \cdot \mathbf{B}$ and allows the scheme to converge to the correct solution. With this scheme, almost every fluid quantity (P, u, ρ, $\mathbf{v}$, and B_y) has converged very well to the exact solution at this resolution. However, in B_x, which should be constant across x in both shocktubes, we see some small, systematic offset from the exact solution still appear at this resolution. This owes to the relatively large $\nabla \cdot \mathbf{B}$ which appears at the discontinuities. Unlike in the Powell scheme, this will converge away, but slowly (because the errors are low-order).

Our CG scheme eliminates this systematic offset. We stress that every other fluid quantity is essentially indistinguishable from the result with the Dedner et al. (2002) scheme, in good agreement with the exact solution. There is still noise/oscillation associated with the discontinuity, but it returns to the correct systematic value. In fact, qualitatively similar noise appears even using a CT scheme; this owes to representing an inclined interface on a mesh (or non-aligned particle configuration), but this is less problematic because it converges away rapidly and does not lead to any systematic errors.

We have also compared the Ryu & Jones (1995) shocktube, and 3D versions of the shocktubes; our conclusions are identical for the same types of discontinuities.

Fig. 2 considers some alternative formulations. First, instead of our CG method, we consider a "locally divergence free" projection as described in § 2.2. As we predicted, this does nothing to reduce $\nabla \cdot \mathbf{B}$ or the systematic errors at the jumps (in fact they are worse), since it does not act on the problematic (non-local) terms. Next, we consider our CG method, but remove the Dedner et al. (2002) divergence-cleaning. While the average behavior is similar to the default CG case, and the systematic error at the shock disappears, we see very large oscillations in the post-shock solution. This is caused by the "overshoots" necessary at faces to obtain a divergence-free reconstruction sourcing additional $\nabla \cdot \mathbf{B}$ waves, which cannot be damped. Finally, we consider the CG method but

---

[7] A public version of FVMHD3D is available at
https://github.com/egaburov/fvmhd3d
[8] This is available at http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html

[9] Here and in all subsequent plots, the CT results are calculated with the grid code ATHENA (Stone et al. 2008), run in its most accurate mode (PPM, CT, CTU). The accuracy and convergence properties of this code are well-studied. All other results are from GIZMO. Further comparison of the codes is found in Hopkins (2015).
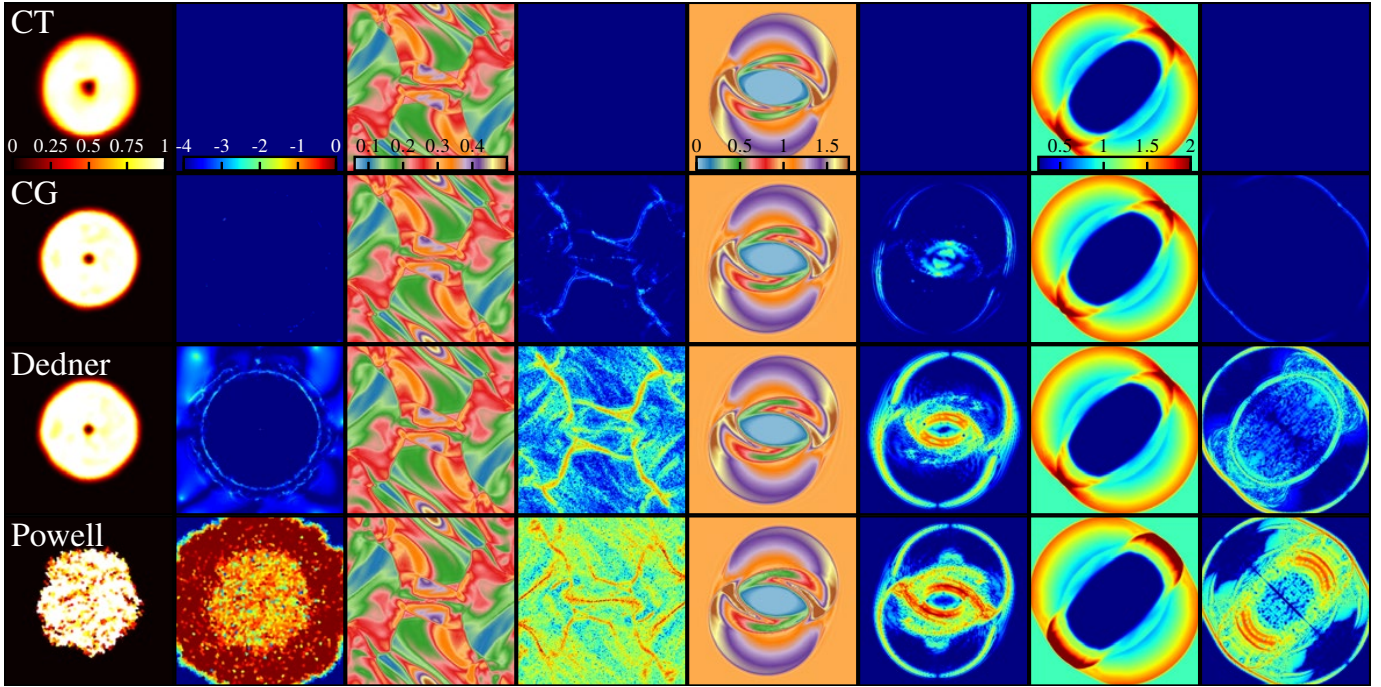
Figure 4. Two-dimensional MHD tests. For each test (column), we compare four methods: CT, CG, Dedner, & Powell (see Fig. 1), as labeled (top-to-bottom). Each pair of columns shows a map of a fluid quantity (*left*), and the corresponding map of $\log_{10}(h_i |\nabla \cdot \mathbf{B}|_i/|\mathbf{B}|_i)$ (*right*), with values following the colorbar. From left to right we show: *Left:* Field loop advection. We show magnetic pressure at time $t = 20$. Methods should preserve a perfect circle at the maximum amplitude; numerical diffusion is visible at the center and edges. *Middle Left:* Orszag-Tang vortex, showing density at $t = 0.5$. *Middle Right:* MHD rotor, showing gas pressure at $t = 0.15$. *Right:* MHD blastwave, showing density at $t = 0.2$. In each case Dedner, CG, and CT solutions are nearly-identical (the extra diffusion in CT in e.g. the field loop, is only because it uses an Eulerian, not Lagrangian code). The Powell scheme produces visibly incorrect features in the blastwave shock jump; and in the field loop test the $h_i |\nabla \cdot \mathbf{B}|_i/|\mathbf{B}|_i$ errors self-interfere and grow unstably. In all cases $h_i |\nabla \cdot \mathbf{B}|_i/|\mathbf{B}|_i$ decreases dramatically from Powell to Dedner and Dedner to CG, where it remains at values $\ll 0.01$.

remove the cleaning and the Powell et al. (1999) terms. Without these terms, even small $\nabla \cdot \mathbf{B}$ errors can still grow unstably, and the system is specifically vulnerable to the tensile instability, which is triggered at the shock. As a result, the oscillations seen before grow unacceptably large. Therefore we will not consider these cases further.

3.2 MHD Waves

While the shocktube problems above illustrate the differences between methods most dramatically, they are less useful as tests of accuracy and convergence (for which we desire smooth problems with known exact solutions). Fig. 3 shows a convergence study in two such problems. First, following Stone et al. (2008), we initialize a traveling fast magnetosonic wave with amplitude $\delta\rho/\rho = 10^{-6}$ (well in the linear regime) in a periodic domain of unit length (with background polytropic $\gamma = 5/3$, density $\rho = 1$, pressure $P = 1/\gamma$, and magnetic field $\mathbf{B}/\sqrt{4\pi} = (1, \sqrt{2}, 1/2)$). After propagating one wavelength, the system should return to its original state so we define the error norm in the perpendicular field as $L1(B_\perp) = N^{-1} \sum_i |B_\perp(x_i, t) - B_\perp(x_i, t = 0)|$. We similarly measure the absolute mean divergence errors. As shown in Paper I, the methods in GIZMO exhibit second-order convergence on this problem. Although our CG method does reduce the magnitude of $h_i |\nabla \cdot \mathbf{B}|_i/|\mathbf{B}|_i$, the effect is only a factor ~ 3 – all methods handle this problem accurately, owing to its relative simplicity and lack of discontinuities. More demanding is the circular polarized Alfven wave test from Tóth (2000); a 3D periodic box of unit size (particle number N_{1D}^3) with $\gamma = 5/3$, $\rho = 1$, $P = 0.1$, $V_\parallel = 0$, $B_\parallel = 1$, $V_\perp = B_\perp = 0.1 \sin(2\pi x_\parallel)$, $V_z = B_z = 0.1 \cos(2\pi x_\parallel)$ (where

$x_\parallel \equiv x \cos\alpha + y \sin\alpha$ and $\tan\alpha = 1/2$ defines the angle of the wave propagation) is evolved until the wave propagates five wavelengths. We then measure the L1 norm in $B_\perp$ and mean $h_i |\nabla \cdot \mathbf{B}|_i/|\mathbf{B}|_i$. This is inherently multi-dimensional and non-linear, making it more demanding. In all cases, there is some numerical wave damping, but again we see second-order convergence. Our CG method again reduces the mean $h_i |\nabla \cdot \mathbf{B}|_i/|\mathbf{B}|_i$ by a similar factor.

In both the linear and non-linear waves, we can see a slight increase in the error norms at low resolution with our CG method. In fact the Powell cleaning actually exhibits the smallest error norms on these problems. Provided the magnetic divergence errors are not so large as to corrupt the solution, this is the "minimal correction" so preserves the accuracy of the underlying method most faithfully. In the CG case, the (small) loss of accuracy owes to the constraint imposed in the reconstruction step – so it cannot always reconstruct the "most accurate" (in a least-squares sense) gradients. This difference appears to converge away relatively quickly at higher resolution.

3.3 Dynamics Test Problems

Figs. 4, 5, & 6 show several 2D tests: advection of a field loop (Gardiner & Stone 2008), the MHD rotor (Balsara & Spicer 1999), the Orszag & Tang (1979) vortex, a strongly magnetized blastwave (Londrillo & Del Zanna 2000), the MHD Rayleigh-Taylor instability (e.g. Jun et al. 1995), and the development of the magneto-rotational instability (MRI) in a shearing-shear simulation (following Guan & Gammie 2008). We show images of fluid quantities and the divergence errors, at some time into the non-linear evolution, and values of fluid quantities. All use 256^2 resolution.
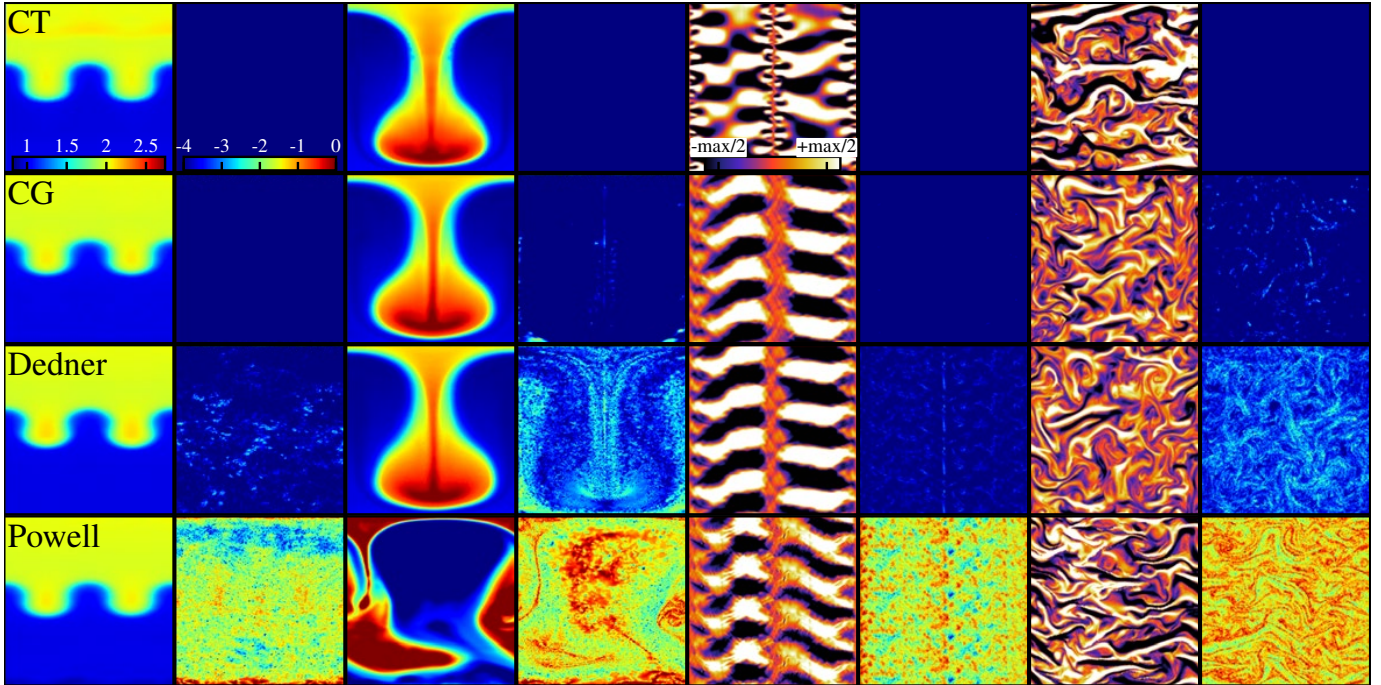
Figure 5. Additional 2D tests, as Fig. 4: MHD Rayleigh-Taylor instability (density is plotted, using a fixed scale given in the left-most colorbar) at times $t = 6$ (*Left*) & $t = 16$ (*Middle Left*), and growth of the magnetorotational instability (MRI) in a shearing sheet at $t = 10$ (*Middle Right*) and $t = 19$ (*Right*). For the MRI we plot the azimuthal/toroidal component B_y of the magnetic field, scaled relative to its maximum absolute value ("max"), so different times can be compared. All methods capture the linear growth of the RT and MRI, and breakup of the non-linear MRI into turbulence at late times. In each, Dedner, CG, and CT agree well (even into non-linear stages); CG maintains $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i \ll 0.01$ even well into non-linear and turbulent evolution. Powell schemes show small deviations in the MRI and linear RT growth, but the errors build up in the non-linear RT and destroy the solution. Note that the different pattern of modes in the CT MRI test owes to different implementations of shearing-box boundary conditions in the grid code used for the CT tests, as compared to the particle code for the other tests shown.

In each case, using the Powell scheme alone leads to *qualitatively* incorrect features (shocks in the wrong position, catastrophic noise, jumps with the wrong shape, etc.), and $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i \sim 1$; these are order-unity errors that do not converge away. In e.g. the RT instability, we see that while the linear (early time) behavior is reasonable, the non-linear behavior is destroyed by growing $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$. In contrast, the Dedner-scheme, CG, and CT results are almost identical at this resolution, in the physical fluid quantities. However the Dedner scheme still produces some small regions where $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|$ can reach values as large as $\sim 0.01 - 0.1$, at this resolution (at sharp discontinuities). These deviations disappear in the CG result, which maintains *maximum* $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i \lesssim 0.01$, even at discontinuities, and mean $\langle h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i \rangle$ about an order of magnitude smaller than the Dedner scheme.

Quantitatively, Fig. 6 illustrates the order-unity errors in shock positions and jumps that appear in the Powell scheme in the rotor and blastwave problems; it also illustrates in the field loop test that the scheme is numerically unstable, and the magnetic energy grows exponentially. The other methods agree well; the linear MRI and RT growth rate are almost identical at this resolution (and agree well with the analytic predictions). Note that in the field loop problem, the CG method is slightly more dissipative compared to Dedner (CT is more dissipative still, but this is primarily due to the CT code being an Eulerian, not Lagrangian, code). This is because there is a real non-zero $\nabla \cdot \mathbf{B}$ set up in the ICs at the "edge" of the circle, as we numerically implement them. This forces the gradients to correct for this, and dissipate away the divergence.

In Fig. 7, we expand our comparison of the magnetic energy in the MRI test to late times. The growth saturates and then the energy

must decay according to the anti-dynamo theorem (in 2D, zero net flux simulations); but the decay rate is sensitive to the numerical diffusivity of the method (see Guan & Gammie 2008, and references therein). We therefore compare our CG, Dedner and Powell results to two different CT implementations in different grid codes. We show this to emphasize that difference between the two CT codes is much larger than between our Dedner and CG results – other factors (beyond the divergence-control scheme) dominate the numerical diffusivity (for a more detailed comparison of methods at different resolution, see Paper I). However, we do see slightly weaker numerical dissipation in our CG compared to (otherwise identical) our Dedner runs. This is surprising given the accuracy comparison in Fig. 3 (we might expect constrained gradients to be slightly less accurate and therefore more diffusive); however, the Dedner terms include explicit wave diffusion and damping sourced by $\nabla \cdot \mathbf{B}$, so the reduction in $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$ in CG leads (non-linearly) to a net reduction in numerical dissipation. More troublingly, in the Powell-only scheme, the large conservation errors that build up lead to a violation of the anti-dynamo theorem and unphysical late-time growth of $|\mathbf{B}|$.

Variations of the above tests with different seed $\mathbf{B}$-fields produce qualitatively identical conclusions. And because the GIZMO methods are Lagrangian, "boosted" or rotated versions of the tests are trivially identical to those shown. We have compared the MHD Kelvin-Helmholtz instability and "blob" test, but the qualitative differences between methods are identical to the RT test shown. The "current sheet" test in Hawley & Stone (1995) is a test of numerical stability which all methods here pass similarly (see Paper I); however as with the field loop we see greater dissipation in CG
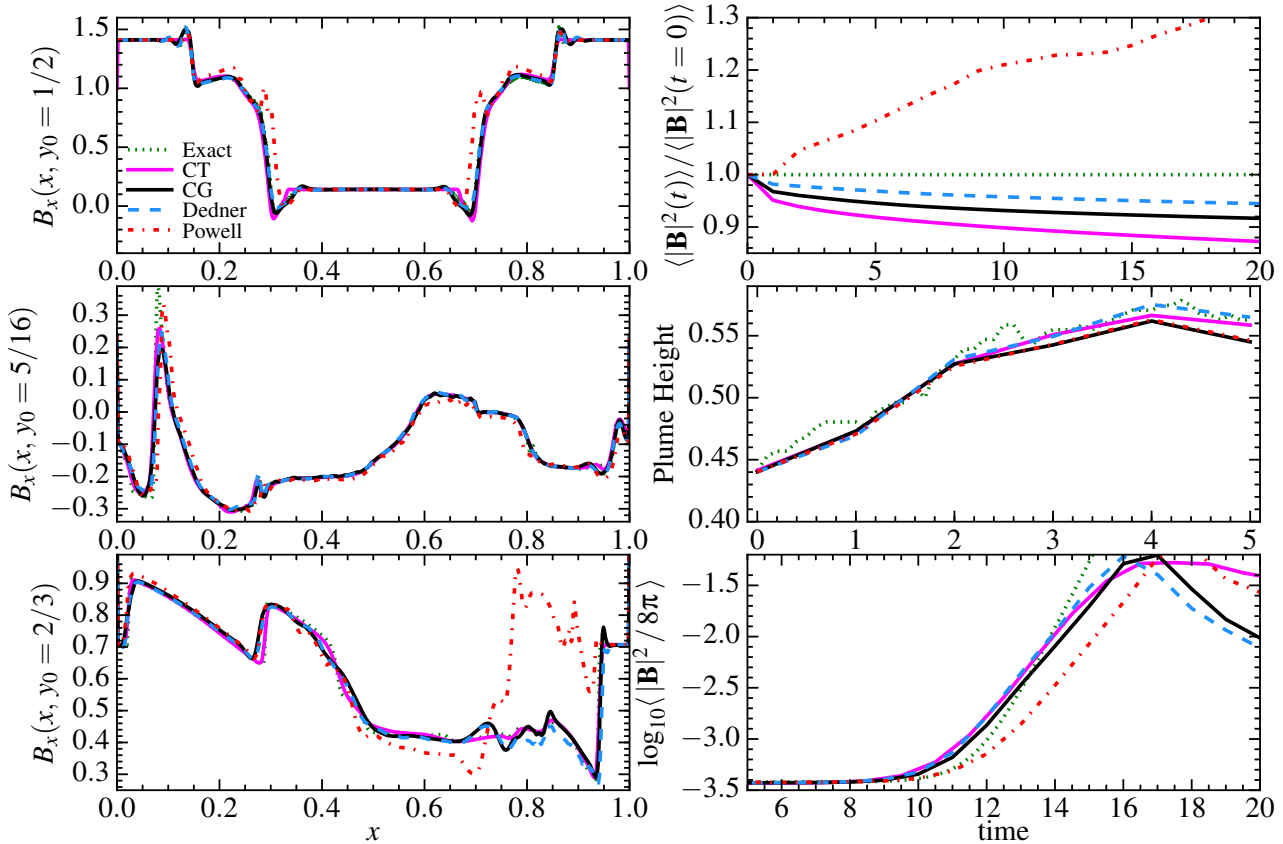
Figure 6. Quantitative comparison of the 2D tests in Figs. 4-5. *Left:* Values of B_x in horizontal slices, for the rotor (*top*), Orszag-Tang vortex (*middle*), and blastwave (*bottom*) tests (at the same time as Fig. 4). We compare methods at 256^2 resolution to an exact solution. All other fluid quantities show comparable or smaller deviations from the exact solution at this resolution. *Right:* Values versus time of the box-averaged $|\mathbf{B}|^2$ in the field loop test (*top*), low-density plume height in the RT test (*middle*), and magnetic energy density in the MRI test (*bottom*). In most tests Powell cleaning produces small deviations (offset shock positions, slower MRI growth); but in the blastwave and field loop tests the failure is dramatic. All other methods agree well and exhibit similar convergence rates. CG shows *slightly* smaller errors at fixed resolution compared to Dedner. Difference between CG & CT (more diffusion for CT in the field loop & RT tests, slightly sharper shock-capturing in CT in the blastwave, & different late-time decay of the MRI) owe to the difference between grid methods (the CT results here) and Lagrangian methods (all others), *not* to divergence errors.

because the ICs contain a real non-zero $\nabla \cdot \mathbf{B}$. We have also simulated low-resolution 3D versions of the RT, KH, field loop, and MRI problems with qualitatively similar results in all cases to the 2D tests here.

As noted by Gardiner & Stone (2008), the field loop problem also provides a useful validation that we are minimizing the "correct" discrete representation of $\nabla \cdot \mathbf{B}$ for our scheme. The problem should at all times have $B_z = \mathbf{B} \cdot \hat{z} = 0$ (where the loop is in the xy-plane); if we initialize the problem with constant but non-zero v_z, then non-zero values of the discrete $\nabla \cdot \mathbf{B} \neq 0$ will produce growth in B_z. In the cleaning-based methods studied here, the Powell et al. (1999) source terms should cancel these errors. Indeed, in each case (Powell, Dedner, and CG), we have run the problem with $v_z = 1$ (comparable to the advection speed in the xy plane) until $t = 20$, and find $B_z \approx 0$ to within machine errors ($\langle |B_z| \rangle < 10^{-20}$) at all times after the first few timesteps. This verifies that the correct discrete source terms for the Powell cleaning are being applied (which are obtained using the *identical* definition of the discrete $\nabla \cdot \mathbf{B}$ used to derive the Dedner and CG terms, in turn). For further validation, we have re-run without the Powell source terms; with no cleaning, $\langle |B_z| \rangle$ grows exponentially (soon exceeding the initial field strengths). With the CG method (but no Powell term), $\langle |B_z| \rangle$ scales (as it should) $\sim h |\nabla \cdot \mathbf{B}|$, jumping to $\sim 10^{-7}$ in a few

timesteps, but remaining around these values as long as we continue running.

3.4 Non-Linear Dynamics with External Forces

Next we consider a set of non-linear dynamics problems where the dominant forces are not MHD, but gravity. These are especially challenging for divergence-control methods because elements are being constantly re-arranged by non-MHD forces in a manner that is often faster than the local fast magnetosonic crossing (hence fluid response) time. Unfortunately they are not rigorous tests, because exact solutions are not known, but they are useful validations that the code does not produce unphysical behaviors or new numerical instabilities under extreme conditions.

We consider (1) collapse of a rotating, self-gravitating protostellar core, to form a protostar and accretion disk which winds up a seed $\mathbf{B}$ field and launch an MHD jet, following Hennebelle & Fromang (2008), (2) the MHD version of the "Santa Barbara cluster," in which a cosmological simulation of gas+dark matter is followed using adiabatic (non-radiative) gas physics, from high redshift until the Lagrangian region being simulated forms an object with the mass of a galaxy cluster at $z = 0$ (Frenk et al. 1999), (3) an isolated (non-cosmological) galaxy disk, with gas, stars, and dark matter, radiative cooling, star formation, and stellar feedback, following the simple sub-grid Springel & Hernquist (2003) "effec-
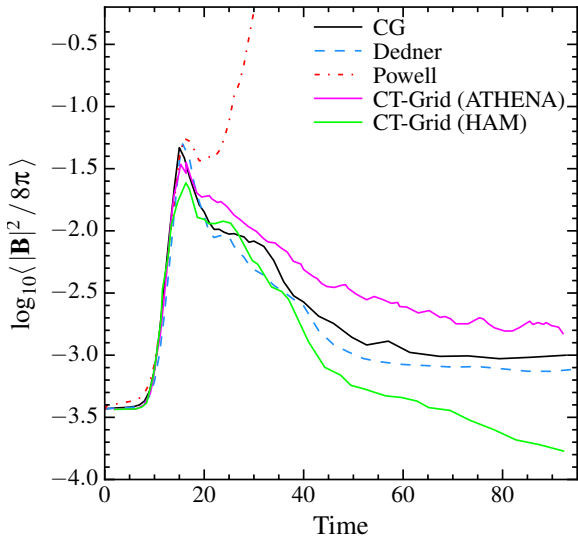
Figure 7. Late-time evolution of the magnetic energy, for the MRI test in Figs. 5-6 (at 256^2 resolution). Physically, the anti-dynamo theorem requires the magnetic energy decay; however the decay *rate* is known to be very sensitive to the numerical dissipation in a given code (greater dissipation producing faster decay). We compare our Powell, Dedner, and CG results in GIZMO to two different CT-based grid codes from Guan & Gammie (2008) (the 2nd-order HAM code and the 3rd-order PPM unsplit CTU result from ATHENA). While the linear growth rates and peak amplitudes are similar, there are significant differences in the decay rate owing to differences in numerical dissipation. Our Dedner and CG results lie between the two CT grid results (with the higher-order CT result the least dissipative, as expected). Interestingly, CG is less dissipative than Dedner alone, even though it may produce a slightly less accurate reconstruction – this owes to the reduction in $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$ producing less explicit dissipation from the Dedner divergence-transport and damping terms. The large errors in conservation from Powell-only cleaning lead to an unphysical non-linear runaway in $|\mathbf{B}|$ (in violation of the anti-dynamo theorem).

tive equation of state" model (in which the phase structure of the ISM is not resolved but replaced with a simple barytropic equation of state, as used in large-volume cosmological simulations; Vogelsberger et al. 2013), and (4) the same disk, treating feedback explicitly according to the FIRE (Feedback in Realistic Environments) project physics (Hopkins et al. 2011, 2012, 2013, 2014; Faucher-Giguere et al. 2015), which explicitly follow the multi-phase ISM, turbulence, and feedback from stellar winds, radiation, and supernovae.

As noted above, these problems do not have known exact solutions, however there are specific qualitative behaviors that should be observed in each case (that must be present owing to basic considerations of conservation and/or dynamics). And they are valuable "stress tests" for divergence control (indeed, many algorithms cannot run these problems without crashing). Fig. 8 summarizes the results. In all cases the behavior is very similar between our Dedner and CG results. In the jet test, we see the jet launched efficiently and the system evolves stably to late times; the mass-loading of these jets is in good agreement with much higher-resolution CT-based AMR results (see Hennebelle & Fromang 2008), as discussed in detail in Hopkins & Raives (2016).[10] Likewise, in the Santa Bar-

---

[10] Unfortunately, we cannot rigorously compare CT methods in Fig. 8, since the relevant physics for these problems is not implemented in the same manner in any CT code. Moreover the gravity solvers are different for these codes, which (since gravity is the dominant force) can introduce much larger differences than the divergence-control method.

bara test, we see the cluster form, and non-linear field amplification in the cluster center; the profiles of density, temperature, magnetic field strength, and velocity also agree well with higher-resolution CT-based AMR runs (compare Miniati & Martin 2011), up to subtle differences at small radii that depend on whether the methods used are Lagrangian or Eulerian (see Hopkins 2015). In the disk problem, the disk remains smooth (as it should) with the "effective EOS" model, while with the FIRE model it rapidly develops super-sonic turbulence, multi-phase structure, and a strong galactic wind. In both cases the field is amplified; amplification is slower in the "effective EOS" case because (by construction) there is no substructure, turbulence, or galactic wind (so only global disk winding amplifies $\mathbf{B}$). In the FIRE case, the molecular gas tracing spiral structure and GMCs is clearly evident in the β map as regions where the thermal pressure is sub-dominant. Although the qualitative behavior is similar, the CG method reduces $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$ by $\sim 1 - 2$ dex, relative to the Dedner case (which, as expected, can reach large $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i \gtrsim 0.1$ in sharp discontinuities).

In all problems here, the Powell-only case is problematic: divergence errors reach $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i \gg 1$, and even $h_i |\nabla \cdot \mathbf{B}|_i / [|\mathbf{B}|_i^2 + 2 P_{\text{thermal}}]^{1/2} \sim 1$. This produces some qualitatively erroneous features: the jet disconnects from its launching zone, and the disk on small scales exhibits a "bending" because the protostar is actually moving with a net z-velocity that grows in time, until it "self-ejects" (owing to the momentum conservation error that comes with subtracting large $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$ terms); this is seen in SPH and grid-based codes using Powell-only control as well (see Price et al. 2012). These should not be possible if the method preserves momentum and problem symmetry accurately. In the cluster and disk problems, the same errors seen in e.g. the field loop problem lead to artificial, unstable growth of magnetic energy: the $\mathbf{B}$ field is higher by about ~ 1 dex at the times shown compared to the Dedner, CG, and CT solutions, and it grows at an unphysical rate (faster than any physical timescales such as the disk dynamical time, even in the smooth-disk case).

We have also compared the 3D MHD Zeldovich pancake test from Paper I with various initial $\mathbf{B}$ (see Zel'dovich 1970; Li et al. 2008). This test, unlike those above, actually has an exact solution and can be compared rigorously. Unfortunately, that is possible because of the highly simplified problem geometry which makes $\nabla \cdot \mathbf{B}$ control unnecessary (even Powell schemes maintain $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i \ll 10^{-4}$), so (not surprisingly), all the methods here produce indistinguishable results from CT.

4 DISCUSSION

We have introduced a new method, Constrained-Gradient MHD, to control the $\nabla \cdot \mathbf{B}$ errors associated with numerical MHD. This involves an iterative approximation to the least-squares minimizing, globally divergence-free reconstruction of the fluid. We implement this in the code GIZMO and show that, compared to state-of-the-art MHD implementations using divergence-cleaning schemes, this is able to further reduce the $\nabla \cdot \mathbf{B}$ errors by orders of magnitude, and improves numerical convergence and accuracy at fixed resolution.

The performance cost of this method is small, compared to constrained transport on irregular grids. It requires one additional neighbor element sweep after the main gradient sweep (the iteration step), in which a the quantity S_0 is re-calculated based on the updated gradients. In our implementation in GIZMO, this increases the CPU cost of the hydro operations by $\sim 20 - 30\%$, with essentially no memory cost.

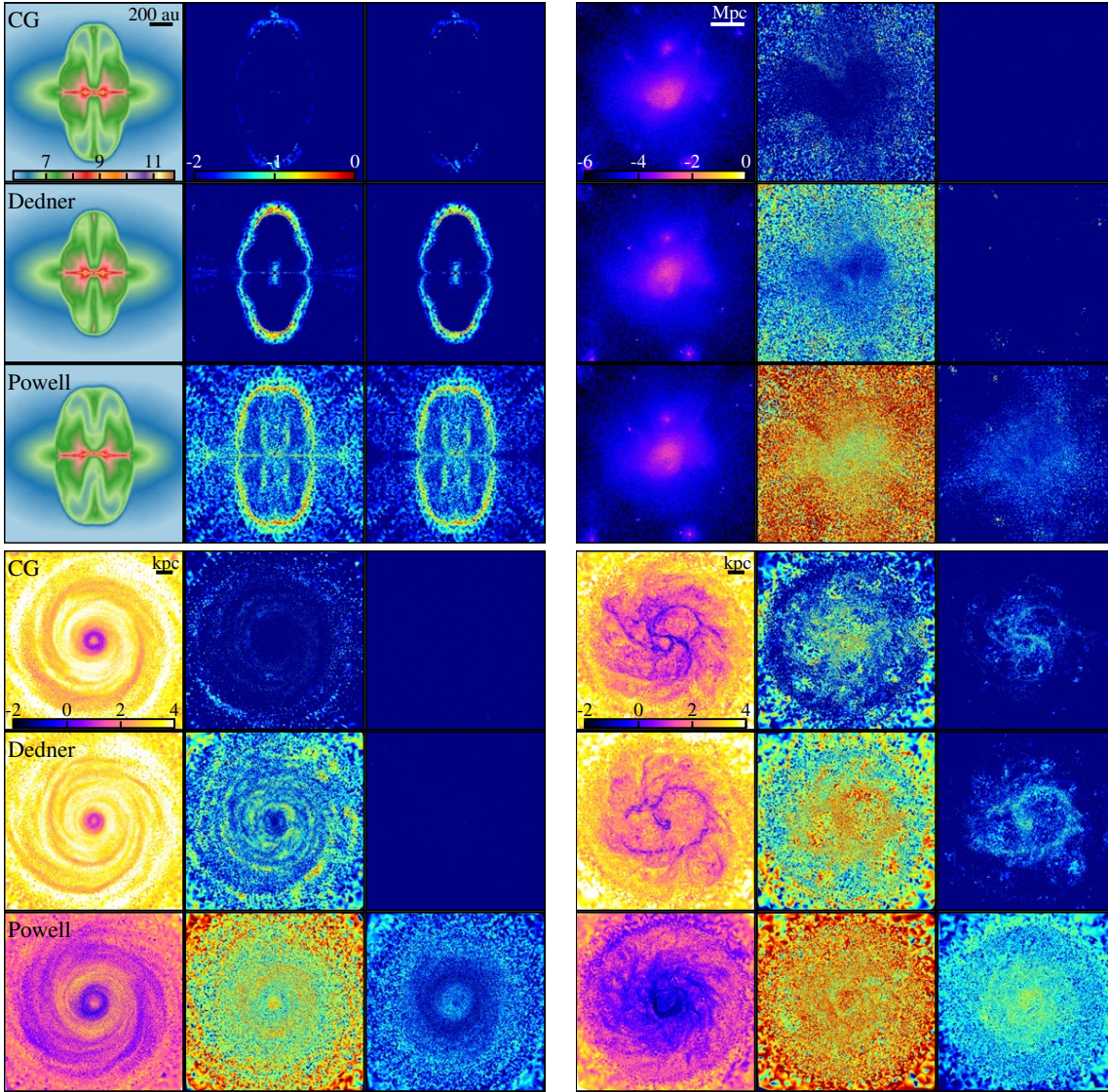This method is motivated in spirit by projection and locally

Figure 8. 3D, non-linear problems with self-gravity. For each we compare CG, Dedner, and Powell methods (rows, as labeled). *Top Left:* Jet formation via collapse of a rotating, magnetized protostellar core (after ≈ 1.1 free-fall times). We plot the density ($\log_{10}[n/\mathrm{cm}^{-3}]$; *left*), divergence errors as before ($h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$; *middle*), and divergence error relative to the total gas pressure ($h_i |\nabla \cdot \mathbf{B}|_i / [|\mathbf{B}|_i^2 + 2 P_{\mathrm{thermal}}]^{1/2}$; *right*), in a slice through the jet axis. A protostar and rotating disk have formed, amplified $\mathbf{B}$, and launched a jet at this stage. *Top Right:* Same quantities, for the magnetized Santa Barbara cluster (a cosmological, non-radiative simulation of dark matter and gas which forms a massive galaxy cluster-hosting halo), at redshift $z = 0$ (slice through the cluster center shown). *Bottom Left:* Isolated star-forming galaxy disk (gas, stars, and dark matter, slice through midplane shown) with radiative cooling and star formation, evolved for $500\,\mathrm{Myr}$, with the smooth, sub-grid "effective equation of state" model for the ISM from Springel & Hernquist (2003). We plot plasma $\beta \equiv P_{\mathrm{magnetic}} / P_{\mathrm{thermal}}$ (*left*), $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$ (*middle*), and $h_i |\nabla \cdot \mathbf{B}|_i / [|\mathbf{B}|_i^2 + 2 P_{\mathrm{thermal}}]^{1/2}$ (*right*). *Bottom Right:* Same quantities for the same disk, but evolved with the Hopkins et al. (2014) FIRE models; these explicitly treat stellar radiation pressure & photo-heating, stellar winds, and SNe, and resolve the multi-phase structure of the ISM and galactic winds. The multi-phase, turbulent structure and constant removal/addition of mass from the system makes this the most challenging case for divergence control. The quantity $h_i |\nabla \cdot \mathbf{B}|_i / [|\mathbf{B}|_i^2 + 2 P_{\mathrm{thermal}}]^{1/2}$ demonstrates that the largest $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$ values correspond to regions where $\mathbf{B}$ is dynamically irrelevant. In all cases, the simulations demonstrate the correct qualitative behaviors; the Dedner cleaning is acceptable, but CG further reduces $|\nabla \cdot \mathbf{B}|$ by $\gtrsim 1$ dex and maintains $h_i |\nabla \cdot \mathbf{B}|_i / [|\mathbf{B}|_i^2 + 2 P_{\mathrm{thermal}}]^{1/2} \lesssim 0.01$. However, the Powell cases are corrupted by large divergence errors: the jet is "puffed out," has detached from the protostellar disk, and the protostar is migrating upwards out of the disk owing to momentum conservation errors associated with large $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$. In the cluster and disk problems, the Powell-only $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i$ reaches > 1, and $\mathbf{B}$-fields are amplified to order-of-magnitude too-large values.

divergence-free reconstruction schemes. However, it avoids the large overhead expense of projection schemes and does not modify the fluxes, and unlike traditional projection schemes it can operate with complicated slope limiters, non-linear gradient estimators, irregular mesh geometries, and adaptive timestepping, and preserves the convergence order of the code. Unlike locally divergence-free schemes, the method here actually minimizes the numerically unstable $\nabla \cdot \mathbf{B}$ terms, as opposed to those from a different estimator.

We consider a large suite of test problems, and compare a variety of divergence-control schemes. With only Powell or "8-wave" cleaning, typical $h_i |\nabla \cdot \mathbf{B}|_i / |\mathbf{B}|_i \sim 1$ in non-linear problems/discontinuities, and this sources zeroth-order systematic errors at discontinuities which do not decrease with increasing resolution. This causes serious problems on a wide range of problems: large systematic errors in shock jumps, catastrophic noise, exponential growth of magnetic energy, and order-unity violations of

energy/momentum conservation and symmetry all appear even at high resolution.

Adding a more sophisticated Dedner hyperbolic-parabolic cleaning-damping term allows the method to converge properly, and reduces $h_i |\nabla \cdot \mathbf{B}|_i/|\mathbf{B}|_i$ by $\sim 1-2$ orders of magnitude. In most tests we find that this is sufficient for ideal behavior and high accuracy; however, non-negligible $\nabla \cdot \mathbf{B}$ errors can still appear at large discontinuities, which lead to small systematic offsets in jump conditions that converge away relatively slowly.

In every case, the CG method further reduces the problematic $\nabla \cdot \mathbf{B}$ terms (by another $\sim 1-3$ orders of magnitude). Critically, while the Dedner scheme requires some finite time and resolution to dissipate non-zero $\nabla \cdot \mathbf{B}$, the CG scheme instantaneously controls $\nabla \cdot \mathbf{B}$ across single resolution elements. As a result, in almost all cases, this means that the *maximum* $h_i |\nabla \cdot \mathbf{B}|_i/|\mathbf{B}|_i < 0.01$, even across arbitrarily sharp discontinuities. This eliminates the systematic offsets present in the Dedner scheme, and gives results which are, for all practical purposes in the tests here, indistinguishable from those obtained using CT methods which maintain $h_i |\nabla \cdot \mathbf{B}|_i/|\mathbf{B}|_i = 0$ to machine precision.

The CG method is easy to implement and quite general: it is applicable to any finite-volume or finite-element method with a well-defined reconstruction procedure (this includes our arbitrary Lagrangian-Eulerian MFM/MFV methods, Godunov and Galerkin-type finite-volume grid/AMR codes, moving-mesh methods, and finite-pointset methods). And it is independent of the "default" procedure used for gradient estimation: all that is needed is a sweep to calculate the terms needed for the correction tensor **G**. Although the formal maintenance of $\nabla \cdot \mathbf{B} = 0$ is not as accurate as CT, it is compatible with arbitrary gradient definitions and slope limiters (these are usually implicitly limited to highly prescribed forms for the magnetic field in CT), is relatively inexpensive, and can trivially handle arbitrary mesh/point geometries, different numbers of spatial dimensions, and adaptive (non-uniform) timesteps.

Finally, we have considered one simple implementation of the method, but it allows considerable freedom which merits further exploration. Different "preferred" gradients (we used a kernel-weighted least-squares estimate), penalty functions, iteration/convergence schemes (e.g. semi-implicit matrix methods), slope-limiters, and cleaning-field (ψ) terms in S_0 could be easily considered and may well provide superior performance. Higher-order generalizations should also be straightforward, although the equations in this paper would need to be modified for a higher-order reconstruction used in determining the $\nabla \cdot \mathbf{B}$ error.

ACKNOWLEDGMENTS

REFERENCES

Balsara, D. S. 2004, ApJS, 151, 149

Balsara, D. S., & Spicer, D. S. 1999, Journal of Computational Physics, 149, 270

Børve, S., Omang, M., & Trulsen, J. 2001, ApJ, 561, 82

Brackbill, J. U., & Barnes, D. C. 1980, Journal of Computational Physics, 35, 426

Brio, M., & Wu, C. C. 1988, Journal of Computational Physics, 75, 400

Dedner, A., Kemm, F., Kröner, D., Munz, C.-D., Schnitzer, T., & Wesenberg, M. 2002, Journal of Computational Physics, 175, 645

Dolag, K., & Stasyszyn, F. 2009, MNRAS, 398, 1678

Evans, C. R., & Hawley, J. F. 1988, ApJ, 332, 659

Faucher-Giguere, C.-A., Hopkins, P. F., Keres, D., Muratov, A. L., Quataert, E., & Murray, N. 2015, MNRAS, 449, 987

Frenk, C. S., et al. 1999, ApJ, 525, 554

Gaburov, E., Johansen, A., & Levin, Y. 2012, ApJ, 758, 103

Gardiner, T. A., & Stone, J. M. 2008, Journal of Computational Physics, 227, 4123

Guan, X., & Gammie, C. F. 2008, ApJS, 174, 145

Hawley, J. F., & Stone, J. M. 1995, Computer Physics Communications, 89, 127

Hennebelle, P., & Fromang, S. 2008, A&A, 477, 9

Hopkins, P. F. 2015, MNRAS, 450, 53

Hopkins, P. F., Cox, T. J., Hernquist, L., Narayanan, D., Hayward, C. C., & Murray, N. 2013, MNRAS, 430, 1901

Hopkins, P. F., Keres, D., Onorbe, J., Faucher-Giguere, C.-A., Quataert, E., Murray, N., & Bullock, J. S. 2014, MNRAS, 445, 581

Hopkins, P. F., Quataert, E., & Murray, N. 2011, MNRAS, 417, 950

—. 2012, MNRAS, 421, 3522

Hopkins, P. F., & Raives, M. J. 2016, MNRAS, 455, 51

Jun, B.-I., Norman, M. L., & Stone, J. M. 1995, ApJ, 453, 332

Kuhnert, J. 2003, in Lecture Notes in Computational Science and Engineering, Vol. 26, Meshfree Methods for Partial Differential Equations, ed. M. Griebel & M. Schweitzer (Springer Berlin Heidelberg), 239–249

Lanson, N., & Vila, J.-P. 2008, SIAM J. Numer. Anal., 46, 1912

Li, S., Li, H., & Cen, R. 2008, ApJS, 174, 1

Londrillo, P., & Del Zanna, L. 2000, ApJ, 530, 508

Luo, H., Baum, J. D., & Löhner, R. 2008, Journal of Computational Physics, 227, 8875

Maron, J. L., & Howes, G. G. 2003, ApJ, 595, 564

Mignone, A., & Tzeferacos, P. 2010, Journal of Computational Physics, 229, 2117

Miniati, F., & Martin, D. F. 2011, ApJS, 195, 5

Mocz, P., Vogelsberger, M., & Hernquist, L. 2014, MNRAS, 442, 43

Monaghan, J. J. 2000, Journal of Computational Physics, 159, 290

Oñate, E., Idelsohn, S., Zienkiewicz, O. C., & Taylor, R. L. 1996, International Journal for Numerical Methods in Engineering, 39, 3839

Orszag, S. A., & Tang, C.-M. 1979, Journal of Fluid Mechanics, 90, 129

Powell, K. G., Roe, P. L., Linde, T. J., Gombosi, T. I., & De Zeeuw, D. L. 1999, Journal of Computational Physics, 154, 284

Price, D. J., & Bate, M. R. 2008, MNRAS, 385, 1820

Price, D. J., & Rosswog, S. 2006, Science, 312, 719

Price, D. J., Tricco, T. S., & Bate, M. R. 2012, MNRAS, 423, L45

Rosswog, S., & Price, D. 2007, MNRAS, 379, 915

Ryu, D., & Jones, T. W. 1995, ApJ, 442, 228

Springel, V. 2005, MNRAS, 364, 1105

Springel, V., & Hernquist, L. 2003, MNRAS, 339, 289

Stone, J. M., Gardiner, T. A., Teuben, P., Hawley, J. F., & Simon, J. B. 2008, ApJS, 178, 137

Swegle, J. W., Hicks, D. L., & Attaway, S. W. 1995, Journal of Computational Physics, 116, 123

Tóth, G. 2000, Journal of Computational Physics, 161, 605

Vogelsberger, M., Genel, S., Sijacki, D., Torrey, P., Springel, V., & Hernquist, L. 2013, MNRAS, in press, arXiv:1305.2913

Yang, H., & Fengyan, L. 2015, ESAIM: Mathematical Modelling and Numerical Analysis

Zel'dovich, Y. B. 1970, A&A, 5, 84
    Stuff