A NEW APPROXIMATE BLOCK FACTORIZATION PRECONDITIONER FOR TWO-DIMENSIONAL INCOMPRESSIBLE (REDUCED) RESISTIVE MHD

ERIC C. CYR†, JOHN N SHADID†, RAYMOND S. TUMINARO†, ROGER P. PAWLowski†, AND LUIS CHACÓN‡

Abstract. The one-fluid visco-resistive MHD model provides a description of the dynamics of a charged fluid under the influence of an electromagnetic field. This model is strongly coupled, highly nonlinear, and characterized by physical mechanisms that span a wide range of interacting time scales. Solutions of this system can include very fast component time scales to slowly varying dynamical time scales that are long relative to the normal modes of the model equations. Fully implicit time stepping is attractive for simulating this type of wide-ranging physical phenomena. However, it is essential that one has effective preconditioning strategies so that the overall fully implicit methodology is both efficient and scalable. In this paper, we propose and explore the performance of several candidate block preconditioners for this system. One of these preconditioners is based on an operator-split approximation. This method reduces the $3 \times 3$ system (momentum, continuity, and magnetics) into two $2 \times 2$ operators: a Navier–Stokes operator (momentum and continuity) and a magnetics-velocity operator (momentum and magnetics) which takes into account the critical Lorentz force coupling. Using previously developed preconditioners for Navier–Stokes, and an initial Schur-complement approximation for the magnetics-velocity system, we show that the split preconditioner is scalable and competitive with other preconditioners, including a fully coupled algebraic multigrid method.

Key words. block preconditioning, magnetohydrodynamics, stabilized finite element, large-scale parallel, multilevel preconditioner, fully implicit

AMS subject classifications. 65F08, 65M60, 76W05

DOI. 10.1137/12088879X

1. Introduction. Our base MHD model is the one-fluid visco-resistive MHD system [23]. This model provides a continuum description of charged fluids in the presence of electromagnetic fields. While this model is generally applicable to both compressible and incompressible flow applications, we focus on a primitive variable formulation in the incompressible limit ($\nabla \cdot \mathbf{u} = 0$) for which approximate block factorization (ABF) and physics-based preconditioners have not been developed. This limit is useful in modeling of various applications such as low-Lundquist-number liquid-metal MHD flows [40, 9], and high-Lundquist-number, large-guide-field fusion plasmas [55, 26, 14]. The B-field formulation of the strong form governing partial differ-
ential equations (PDEs) written as residuals is

\begin{equation}
R_u = \rho \frac{\partial u}{\partial t} + \rho (u \cdot \nabla u) - \nabla \cdot (T + T_M),
\end{equation}
\begin{equation}
R_p = \rho \nabla \cdot u,
\end{equation}
\begin{equation}
R_B = \frac{\partial B}{\partial t} - \nabla \times (u \times B) + \nabla \times \left( \frac{\eta}{\mu_0} \nabla \times B \right),
\end{equation}

where the viscous and magnetic stress tensors are

\begin{equation}
T = -p I + \mu [\nabla u + \nabla u^T],
\end{equation}
\begin{equation}
T_M = \frac{1}{\mu_0} B \otimes B - \frac{1}{2\mu_0} |B|^2 I.
\end{equation}

The dependent variables, velocity \( u \), hydrodynamic pressure \( p \), and the magnetic field \( B \), satisfy the momentum equation (1.1), the continuity equation (1.2), and the magnetics evolution equation (1.3) when \( R_u = R_B = 0 \) and \( R_p = 0 \). The transport properties, \( \rho, \mu, \eta, \mu_0 \), are the density, dynamic viscosity, magnetic resistivity, and the magnetic permeability of free space, respectively. Constitutive equations define the Newtonian stress tensor \( T \) in (1.4). Ampere’s law, neglecting the displacement current, provides the plasma current, \( J = 1/\mu_0 \nabla \times B \), and the magnetic stress tensor \( T_M \) is obtained by representing the Lorentz force term \( (J \times B) \) as the divergence of a tensor.

For the purposes of this study we focus on a two-dimensional (2D) geometry, as in our previous work on scalable fully coupled algebraic multigrid (AMG) preconditioners for primitive variable incompressible resistive MHD [50]. This choice enables the study of the essential characteristics of the impact of the coupling of the momentum and magnetic field equations and allows the analysis of efficient preconditioning methods for this system. In 2D, the in-plane magnetic field can be expressed in terms of the third component of the vector potential \( A_z \) as \( B = \nabla A_z \times z \). The governing equation for \( A_z \) is then given by [6]:

\begin{equation}
R_A = \frac{\partial A_z}{\partial t} + u \cdot \nabla A_z - \frac{\eta}{\mu_0} \nabla^2 A_z + E_z^0.
\end{equation}

Similarly, the magnetic stress can be rewritten in terms of \( A_z \) as

\[ T_M = \frac{3}{2\mu_0} \left[ \frac{\partial A_z}{\partial y} \right]^2 (\hat{e}_x \otimes \hat{e}_x) + \frac{1}{\mu_0} \left[ \frac{\partial A_z}{\partial y} \frac{\partial A_z}{\partial x} \right] (\hat{e}_x \otimes \hat{e}_y) - \frac{1}{\mu_0} \left[ \frac{\partial A_z}{\partial y} \frac{\partial A_z}{\partial x} \right] (\hat{e}_y \otimes \hat{e}_x) + \frac{3}{2\mu_0} \left[ \frac{\partial A_z}{\partial x} \right]^2 (\hat{e}_y \otimes \hat{e}_y). \]

Here the unit direction vectors \( \hat{e}_x, \hat{e}_y, \hat{e}_z \) and the tensor product identify the components of the magnetic stress tensor. After this transformation, the dependent variables \( u, p, A_z \) satisfy (1.1), (1.2), and (1.6) when \( R_u = 0 \) and \( R_p = R_A = 0 \).

The governing balance equations are discretized in space using a stabilized finite element method that is reviewed below. The time discretization is based on a method of lines approach and employs implicit time integration algorithms. Details of these implementations can be found in [50]. Our approach is to solve the resulting fully coupled system of nonlinear algebraic equations using a Newton–Krylov (NK) method [11]. However, due to the the elliptic incompressibility constraint (1.2) and the Alfvén wave, the sequence of discrete linear systems are typically ill-conditioned.
We emphasize that this difficulty arises not from the specific aspects of the stabilized finite element discretization, but instead is a result of the MHD physics itself. To alleviate this difficulty, and achieve an efficient and scalable NK method, an effective preconditioner for these linear operators is required.

One approach to preconditioning that has demonstrated significant success in large-scale multiphysics applications [56, 53, 37, 38], including this MHD formulation [50], uses a fully-coupled aggressive coarsening AMG algorithm. This method relies on all the unknowns being represented at each of the finite element nodes and constructs projections of the entire fully coupled system. Another approach to designing preconditioners for multiphysics systems exploits a decomposition of the physics into subsystems where multigrid provides an effective approximation of the action of the inverse. A common technique to determine candidate decompositions is to carry out an analysis of the strength of interactions of the component PDE mechanisms to identify strong off-diagonal block coupling [31]. Related techniques also employ and extend existing operator-split and semi-implicit time integration solution methods as preconditioners, since these often encode the strong-coupling mechanisms to produce viable reduced-coupling solution methods [33]. These “physics-based” preconditioners have been applied successfully to the shallow water equations [31, 41], radiation-diffusion systems [42], fluid flow with phase change [34], and several MHD formulations [6, 5, 3].

The approach pursued here is closely related to the physics-based approach in that it decomposes the system into its physical components, each of which is amenable to AMG. However, our approach differs in that we first consider an ABF of the discrete Jacobian matrix that is typically motivated by forming an approximate block LU or LDU factorization. This approach proceeds by approximating the needed block inverses with AMG. Effective preconditioners must carefully consider the spectral properties of the component block operators and the approximate Schur complement operators which account for coupling between the physical components (for MHD this is velocity, pressure, and the vector potential). Through this linear algebraic view of preconditioning, a simplified system of block component equations is developed that encodes a specific physics-based decomposition for many applications. Therefore, these methods have a direct correspondence to the physics-based methods described above. ABF preconditioners have been explored in considerable depth for Navier–Stokes, and have been shown in prior studies to lead to scalable linear solvers [16, 17, 39, 18, 20, 8].

In this study, we propose three ABFs for our MHD formulation. The first factorization preconditioner uses an upper diagonal approximation for the preconditioner that includes a unidirectional coupling from magnetics to momentum, but leaves the fluid solve fully coupled. The second preconditioner was briefly presented in [52] without analysis or detailed study. This preconditioner uses an operator-split approach where the Navier–Stokes coupling derived from the incompressibility condition is handled separately from the coupling that supports the Alfvén wave. This preconditioner is very similar to one recently developed, in a concurrent effort, for the Navier–Stokes equations employing a dimensional splitting [1]. The final factorization uses a Neumann series expansion to construct approximate Schur complement operators.

This work differs from previous work in physics-based preconditioning for resistive MHD in that we are using a primal fluid-variable (velocity-pressure) incompressible MHD formulation. The work in [5] explores preconditioning of an incompressible sys-

---

1 For example, this is the case for equal-order stabilized finite element discretizations.
tem using a current-vorticity-stream function formulation of MHD. This formulation simplifies the structure of the linear system by implicitly satisfying the incompressibility constraint. The primary distinction from this work is that we need to explicitly handle the Navier–Stokes coupling derived from the incompressibility constraint. Additionally, a compressible MHD formulation was considered in [3]. As we will see, prior work in compressible MHD is not directly applicable here because incompressible MHD lacks a density time derivative that can considerably simplify the Schur complement approximation. This seemingly small difference allows block factorizations that are not possible for the incompressible case.

The remainder of this paper is organized as follows. In section 2, we describe the stabilized finite element method that is used to discretize the MHD equations. In addition, the component block form of the Jacobian operator from which the block preconditioners are derived is presented. Section 3 presents three new ABF preconditioners for MHD. For one of these preconditioners, the split ABF method, we show that the error in the split approximation is in some sense small. A subsequent analysis of the spectrum of the preconditioned operator indicates that this preconditioner may lead to good convergence in the context of Krylov solvers. Results demonstrating the scalability of the split preconditioner are presented in section 4, with a comparison against the other ABF preconditioners and existing fully coupled algebraic preconditioners including algebraic multigrid. Finally, some conclusions are drawn in section 5.

2. Stabilized finite element formulation. Equations (1.1), (1.2), and (1.6) present the governing equations in convected form, for momentum, total mass, and vector potential in residual notation, respectively. We approximate this continuous PDE problem using an equal-order stabilized finite element formulation. This stabilized finite element method avoids stability and algorithmic limitations of mixed Galerkin finite element formulations. In particular mixed Galerkin finite element formulations of the momentum-continuity equations of the Navier–Stokes part of the MHD system must satisfy the Ladyzhenskaya–Babuska–Brezzi (LBB) stability condition (see, e.g., [24]). This condition prevents the use of equal-order finite element spaces defined with respect to the same finite element mesh. A final difficulty is that for highly convected flows, solved with coarse to moderate resolution discretizations, an oscillatory instability can be generated with Galerkin methods.

Consistently stabilized finite element methods for Navier–Stokes address the issues described above by using a combination of properly weighted residuals of the governing balance equations. These methods simultaneously relax the incompressibility constraint, and add streamline diffusion to the weak equations to limit oscillations in highly convected flows [13]. The specific stabilized finite element formulations employed in this study are shown in Table 2.1. The intrinsic-time-scale stability parameters (\( \hat{\tau}_m \) and \( \hat{\tau}_{A_z} \)) are based on the formulations of Hughes and Mallet [29] and Shakib [54] for Navier–Stokes, with an adaptation of the stabilized formulation of Codina and Hernandez-Silva [7] for resistive MHD. Details of this formulation can be found in the appendix. More details on the particular stabilized formulation for MHD can be found in [50]. It should be noted however that the ABF preconditioners that are presented in this paper are generally applicable to both Galerkin and stabilized discretizations of the resistive MHD system we consider.

The nonlinear weak formulation in Table 2.1 is solved with an inexact Newton method [15] that has been demonstrated to be robust for CFD and MHD systems [53, 50]. The discrete form of the matrix equations that results from the Newton
Table 2.1

Stabilized finite element formulation of transport/reaction PDEs, where the residual equations \( R_i \) are presented in (1.1), (1.2), and (1.6). Here \( \Phi \) is a global weighting function used to formally define the weak form. The sum \( \sum_e \) indicates the integrals are taken only over element interiors \( \Omega_e \) and integration by parts is not performed.

<table>
<thead>
<tr>
<th>Table 2.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Momentum</td>
</tr>
<tr>
<td>( F_{u,i} = \int_{\Omega} \Phi R_{u,i} , d\Omega + \sum_e \int_{\Omega_e} \rho \mathbf{\hat{r}}<em>m (u \cdot \nabla \Phi) R</em>{u,i} , d\Omega )</td>
</tr>
<tr>
<td>Total mass</td>
</tr>
<tr>
<td>( F_P = \int_{\Omega} \Phi R_P , d\Omega + \sum_e \int_{\Omega_e} \rho \mathbf{\hat{r}}_m \nabla \Phi \cdot \mathbf{R} , d\Omega )</td>
</tr>
<tr>
<td>Z-vector potential</td>
</tr>
<tr>
<td>( F_{A_z} = \int_{\Omega} \Phi R_{A_z} , d\Omega + \sum_e \int_{\Omega_e} \hat{A}<em>z (u \cdot \nabla \Phi) R</em>{A_z} , d\Omega )</td>
</tr>
</tbody>
</table>

Linearized stabilized finite element discretization of the governing balance PDEs is

\[
\begin{bmatrix}
F & B^T & Z
\end{bmatrix}
\begin{bmatrix}
\Delta u \\
\Delta p \\
\Delta A_z
\end{bmatrix}
= \begin{bmatrix}
-r_u \\
-r_p \\
-r_{A_z}
\end{bmatrix}.
\]

In this representation, the vectors, \( \Delta u, \Delta p, \Delta A_z \), contain the Newton updates to the nodal velocities, pressures, and vector potential respectively. The block matrix \( F \) corresponds to the combined discrete transient, convection, diffusion, and stress terms acting on the unknowns \( \Delta u, B^T \) corresponds to the discrete gradient operator; \( Z \) is the Lorentz force operator; \( B \) is the divergence operator; \( C \) corresponds to the discrete “pressure Laplacian” type operator that is generated by the pressure stabilization [13]; \( Y \) is a vector mass-matrix-type operator scaled by the gradient components of \( A_z \); and \( D \) is a combined discrete transient, convection, diffusion operator acting on \( \Delta A_z \). The vectors \( r_u, r_p, \) and \( r_{A_z} \) contain the right-hand side residuals for Newton’s method. The existence of the weak form Laplacian matrix, \( C \), in the stabilized finite element discretization is in contrast to Galerkin methods using mixed interpolation that produce a zero block on the diagonal of the total mass equation. The existence of the block matrix \( C \) helps to enable the fully coupled solution of the linear systems with a number of algebraic and domain decomposition type preconditioners that rely on nonpivoting ILU type factorization, or in some cases methods such as Jacobi or Gauss–Seidel as subdomain solvers [51, 50].

3. Block preconditioners. Preconditioners can often be constructed using block factorization ideas; see [17] and the references therein. In this paper, we focus on the following block LU decomposition:

\[
(3.1) \quad \mathcal{J} = \begin{bmatrix}
F & B^T & Z \\
B & C & 0 \\
Y & 0 & D
\end{bmatrix} = \begin{bmatrix}
I & -B F^{-1} I & -B F^{-1} Z \\
B F^{-1} I & S & -B F^{-1} Z \\
Y F^{-1} & -Y F^{-1} B^T S^{-1} & I
\end{bmatrix} \begin{bmatrix}
F & B^T & Z \\
S & -B F^{-1} Z \\
P
\end{bmatrix},
\]

where \( S \) is the fluid Schur complement for Navier–Stokes given by

\[
(3.2) \quad S = C - B F^{-1} B^T
\]

and

\[
(3.3) \quad P = D - Y F^{-1} (I + B^T S^{-1} B F^{-1}) Z
\]

is a Schur complement coupling velocity, pressure, and magnetics. Those familiar with block factorization techniques for incompressible Navier–Stokes will notice that (3.1)
is significantly more complex than corresponding factorizations for $2 \times 2$ block systems. In particular, the definition of $P$ includes the operator $S^{-1}$. This nesting of Schur complements makes the efficient approximation of $P$ quite challenging, not to mention an approximation for $P^{-1}$ which is typically required within a factorization-based preconditioner. It should be noted that alternative equation orderings are possible and these lead to different factorizations and forms of the Schur complement. For example, a $[p, A_z, u]$ ordering is considered in [3] for compressible MHD. In this case, $C$ is a transient convection-diffusion operator for density (which for small $\Delta t$ can be effectively approximated by a scaled mass matrix) and because of the arrowhead structure the Jacobian is essentially $2 \times 2$ and therefore the Schur complement is not nested. For incompressible MHD, this ordering is not useful as the stabilization operator, $C$, is not necessarily invertible. Other orderings, such as $[u, A_z, p]$, are also possible. However, these lead to nested Schur complements that are also difficult to approximate. The ordering used here was chosen because it leaves the fluid Schur complement intact and will permit application of existing Navier–Stokes preconditioning technology to approximate $S^{-1}$.

One straightforward preconditioner is based on ignoring the contribution of $Y$, which leads to the upper triangular matrix

$$(3.4) \quad M_{BikUp} = \begin{bmatrix} F & B^T & Z \\ B & C & D \end{bmatrix}. $$

The justification for dropping $Y$ is based on observing that $Y$ is primarily a vector mass-matrix operator. Thus, its application to a vector $u$ does not correspond to differentiation of $u$. As $Y$ contains no derivative terms it might then seem less important than the other terms (e.g., $Z$). To approximate the action of $M_{BikUp}$, a Navier–Stokes preconditioner can be applied to the upper $2 \times 2$ block, and a simple multigrid solver can be applied to approximate $D^{-1}$. This preconditioner successfully handles the elliptic incompressibility constraint. However, it ignores the stiffness generated by the coupling of the momentum and magnetics equation. As will be demonstrated, capturing the stiffness associated with the Alfvén wave (represented by the $P$ Schur complement) is critical to developing an effective MHD preconditioner over a broad range of problems, including those with strong and weak coupling of the fluid flow and electromagnetic effects.

Given the difficulty of approximating the nested Schur complement $P$, we consider an alternate ABF preconditioner. The basic idea is motivated from operator-splitting solution methods for the coupled system (see [33, 46] for other physics-based approaches). In particular, we treat the fluid flow and magnetics-velocity systems as independent $2 \times 2$ operators by considering an approximation to the Jacobian, $J$, and its corresponding factorization:

$$(3.5) \quad M_{Split} = \begin{bmatrix} F & B^T & Z \\ B & C & D \\ Y & YF^{-1}B^T & D \end{bmatrix} = \begin{bmatrix} F & Z \\ I & I \\ Y & D \end{bmatrix} \begin{bmatrix} F^{-1} & I \\ I & I \\ B & C \end{bmatrix}.$$

Notice that the leftmost and rightmost operator in the factorization have one row and column containing only the identity on the diagonal; thus the inverse simplifies to approximating the inverse of a block $2 \times 2$ operator. To emphasize this simplification, we refer to these two operators as reduced $2 \times 2$. The rightmost reduced $2 \times 2$
matrix ignores magnetics terms and thus reduces this subsolve to an incompressible fluid flow problem. Several possible Navier–Stokes preconditioners can now be considered to approximate its inverse \[16, 17, 39, 18, 20, 8, 1\]. The leftmost reduced \(2 \times 2\) matrix ignores the incompressibility constraint and corresponds to a magnetics-velocity system. This magnetics-velocity system, to our knowledge, has not been considered before. We consider several approximations to its inverse as discussed in section 3.2.2 that are based on Navier–Stokes ideas and compressible MHD preconditioners \[5, 4\]. We will see that (3.5) avoids the difficulties of nested Schur complements while properly addressing the stiffness from both the incompressibility constraint and the Alfvén wave if appropriate approximations are made for these reduced \(2 \times 2\) inverses. The boxed term in the unfactored expression for \(M_{\text{Split}}\) highlights the error made in approximating \(J\) with \(M_{\text{Split}}\). Ultimately, it is the nature of this term that determines whether or not \(M_{\text{Split}}\) is an appropriate preconditioner for \(J\). Interestingly, this error is a structurally small perturbation of the original operator, though this does not imply that the total effect of this error is small.

A similar preconditioner motivated by the results of Murphy, Golub, and Wathen \[43\] for Stokes flow systems is developed using only the upper triangular factor of an LU factorization of the magnetics-velocity operator in (3.5). This gives rise to

\[
M_{\text{Split}-r} = \begin{bmatrix} I & Z \\ I & \hat{P} \end{bmatrix} \begin{bmatrix} F & B^T \\ B & C \end{bmatrix} I,
\]

where

\[
\hat{P} = D - YF^{-1}Z
\]

is the Schur complement derived from the magnetics-velocity \(2 \times 2\) operator. The action of a preconditioner based on (3.6) requires one less application of \(F^{-1}\) and is often less expensive than preconditioners associated with (3.5); hence we refer to \(M_{\text{Split}-r}\) as reduced.

3.1. Analysis of the spectrum. The convergence of a GMRES solver is strongly influenced by the spectrum of the operator. For a right preconditioned version of the split preconditioner, it is easy to show that

\[
JM_{\text{Split}}^{-1} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ K_u - (D - K_u Z)P^{-1}YF^{-1} & K_p & (I - YF^{-1}B^TS^{-1}BF^{-1}Z\hat{P}^{-1}) \end{bmatrix},
\]

where

\[
[K_u \ K_p] = \begin{bmatrix} Y & 0 \\ B & C \end{bmatrix}^{-1} = YF^{-1} \begin{bmatrix} I + B^TS^{-1}BF^{-1} & -B^TS^{-1} \end{bmatrix}.
\]

The operator on the lower right diagonal is a (not necessarily small) perturbation of the identity

\[
I - YF^{-1}B^TS^{-1}BF^{-1}Z\hat{P}^{-1}
\]

\[
= \left(\hat{P} - YF^{-1}B^TS^{-1}BF^{-1}Z\right)\hat{P}^{-1}
\]

\[
= \left(D - YF^{-1}(I + B^TS^{-1}BF^{-1})Z\right)\hat{P}^{-1}
\]

\[
= P\hat{P}^{-1},
\]
where $P$ is the nested Schur complement defined in (3.3) and $\hat{P}$ is defined in (3.7). Thus, the spectrum of the preconditioned operator is

$$\lambda(JM_{\text{Split}}^{-1}) \subseteq \lambda(P\hat{P}^{-1}) \cup \{1\},$$

where $P\hat{P}^{-1}$ is equivalent to preconditioning the nested Schur complement with a Schur complement derived from the $2 \times 2$ magnetics-velocity system. Analogously, it is possible to show that a right preconditioned system based on (3.6) is

$$JM_{\text{Split}}^{-1} - r = \begin{bmatrix}
I & 0 & 0 \\
0 & I & 0 \\
K_u & K_p & P\hat{P}^{-1}
\end{bmatrix}.$$

Thus, we expect the convergence behavior of $M_{\text{Split}}$ and $M_{\text{Split}}^{-1} - r$ to be similar as the spectrums of the two preconditioned operators are equivalent.

Some further analysis of $P\hat{P}^{-1}$ is possible when $C = 0$ (as is common with a mixed Galerkin finite element formulation). In particular, the operator

$$\Pi = I + B^T S^{-1} B F^{-1}$$

is an idempotent projection. Thus, it annihilates vectors in the range of $B^T$, i.e., $\Pi B^T = 0$ while it does not alter vectors $y = Fs$ if $s$ lies in the null space of $B$. As $B^T$ represents a gradient operator, $\Pi$ generally annihilates gradients of scalar functions. Similarly, $B$ represents a divergence operator and so $\Pi Fs \approx Fs$ when $s$ is a curl of a vector field. This follows from the vector identity $\nabla \cdot (\nabla \times A) = 0$.

Returning to $P\hat{P}^{-1}$, we see that its eigenvalues are equivalent to those of

$$\hat{T} = \hat{P}^{-1} (D - YF^{-1}\Pi Z)$$

which is obtained by applying a similarity transformation to $P\hat{P}^{-1}$. Thus, we now investigate the eigenvalues of $\hat{T}$. Given the above idempotent discussion, it follows that if a vector $v$ exists such that $v$ is an eigenvector of $\hat{T}$ and $Zv$ is an eigenvector of $\Pi$, then the corresponding eigenvalue of $\hat{T}$ is either unity (when $Zv$ is associated with an eigenvalue of unity for $\Pi$) or identical to an eigenvalue of $P\hat{P}^{-1} D$ (when $Zv$ is associated with a zero eigenvalue of $\Pi$). In general, the eigenvectors of $\hat{T}$ do not give rise to $Zv$ which lie totally in the null space of $\Pi$ or its complement and so the above argument is certainly oversimplified. The remainder of this subsection, however, illustrates that on a model problem the eigenvalues of $\hat{T}$ roughly lie somewhere between those of $P^{-1} D$ and one.

As a complete analysis of $P\hat{P}^{-1}$ for the stabilized finite element formulation of the general PDE model is difficult, we instead pursue a Fourier analysis for a simplified discretization and PDE model that preserves the essential coupling of the vector potential MHD system. To this end we consider a model PDE (constant velocity flow) with a marker-and-cell discretization, where the pressure and vector potential are located at cell centers and velocities are defined on cell faces [25]. The boundary conditions are periodic and the mesh is an $N \times N$ uniform grid such that Fourier analysis can be applied. All operators are given by constant coefficient stencils, and the linearization is chosen so that stiff modes, the elliptic incompressibility constraint, and the Alfvén wave are represented. In particular, we consider

$$F \sim \rho \Delta t^{-1} + p u \cdot \nabla - \mu \nabla \cdot \nabla,$$

$$Y \sim \nabla A_z \cdot,$$

$$Z \sim 1/\mu_0 \nabla A_z \nabla \cdot \nabla,$$

$$D \sim \Delta t^{-1} + u \cdot \nabla - \eta/\mu_0 \nabla \cdot \nabla,$$

$$}
where \( u = [0; 1], \nabla A_z = [c_1; c_2], \) and \( \Delta t \) is the implicit Euler time step. The operators \( B \) and \( B^T \) correspond to the usual divergence and gradient, and \( C \) for this stable discretization is zero. This precise choice of both the PDE and the discretization allows the operator \( PP^{-1} \) to be diagonalized by the Fourier transform. That is, define

\[
W_{kj} = \frac{e^{2\pi i j_2 k_1/N} e^{2\pi i j_2 k_2/N}}{\beta_j},
\]

where

\[
k_1 = 0, \ldots, N, \, \, k_2 = 0, \ldots, N, \, \, j_1 = 0, \ldots, N, \, \, j_2 = 0, \ldots, N,
\]

\[
k = k_2 N + k_1 + 1, \, \, j = j_2 N + j_1 + 1,
\]

and \( \beta_j \) is chosen so that \( ||W_{\cdot j}||_2 = 1 \). Then, \( W^T P \hat{P}^{-1} W = \Lambda \) and \( \Lambda \) is a diagonal matrix with the eigenvalues of \( PP^{-1} \) as entries. Additionally, \( W \) diagonalizes \( P, D, \) and \( \hat{P} \). It should be noted that there is a corresponding \( 2 \times 2 \) block diagonal matrix with \( W \) for each block diagonal entry, denoted by \( \hat{W} \), which diagonalizes \( F \). Additionally, \( W^T \hat{Z} W, W^T B^T \hat{W}, \) and \( \hat{W}^T Y^T W \) are essentially all bidiagonal matrices with two nonzero per column. Analytic expressions for the eigenvalues of \( \hat{P}^{-1}, \hat{D}^{-1}, \) and \( PD^{-1} \) have been developed involving trigonometric functions. As these are relatively complicated, we omit them and instead consider a visual representation of the eigenvalue spectrum.

Figure 3.1 presents a series of nine images depicting the eigenvalues of \( \hat{D}^{-1} \), \( PP^{-1} \), and \( PD^{-1} = (PD^{-1})^{-1} \) for different time steps and for different scalings of the components in \( Y \) and \( Z \) (i.e., different \( c_1 \) and \( c_2 \)). The physical parameters are equivalent to those used for the island coalescence problem discussed in section 4.3. In this problem regime, the Alfvén wave plays an important role in the solution. Notice that we plot the inverse of \( PD^{-1} \). As the condition number of a matrix and its inverse are equal, the plot of the inverse matrix is equally relevant for comparing condition numbers, and in this case it is easier to compare with the other plotted values. The time step increases from the top row of plots to the bottom. The direction of the magnetic field, defined by \( B = \nabla \times [0; 0; A_z] \), varies left to right across the plots. Overall, one can see that for the largest time steps, the spread of eigenvalues is much greater for \( DP^{-1} \) and so we would expect this to be a significantly poorer preconditioner than the split preconditioner \( PP^{-1} \). Further, some eigenvalues of \( DP^{-1} \) extend into the left half-plane which can be problematic for iterative solvers. Generally, however, the difference between the eigenvalues for \( PP^{-1} \) and \( DP^{-1} \) diminishes noticeably for small time steps. Thus, one would expect comparable convergence rates for small time steps. This basic behavior and sensitivity to the time step will be confirmed by numerical experiments in the results section.

Finally, one can observe that the eigenvalues of \( PP^{-1} \) appear to lie somewhere in between those of \( \hat{D}^{-1} \) and one. This is most evident for the leftmost figures where one can clearly see that the blue pattern mirrors a version of the red pattern that is \textit{pulled} toward the value one. A similar relationship is apparent in the other images when one examines them closely (though this is not shown here). This provides some support to the notion that eigenvalues of \( T \) lie between \( \hat{D}^{-1} \) and one. Intuitively, it follows that the Schur complement approximation obtained from the split preconditioner should be superior to using \( D \) to approximate \( P \). In particular, the eigenvalue spectrum associated with the split preconditioner is somewhat \textit{better}
than that of $\tilde{D}\tilde{P}^{-1}$ as it is pulled toward the value one. At the same time, we expect (as illustrated) that the eigenvalue spectrum of $D\tilde{P}^{-1}$ is better than the spectrum of $DP^{-1}$ due to the fact that $\tilde{P}$ is a standard Schur complement directly involving $D$ while $P$ is a nested Schur complement which also includes incompressibility that is not accounted for by the operator $D$.

3.2. MHD split algorithm. The preconditioner given by (3.5) can be implemented by the following two step algorithm:

$$\hat{x} = \text{SplitPrec}(\mathcal{J}, b):$$

1) $x^* = \begin{bmatrix} F & I & Z \end{bmatrix}^{-1} b,$ 2) $\hat{x} = x^* + \begin{bmatrix} F & B^T \\ B & C \\ I \end{bmatrix}^{-1} (b - \mathcal{J}x^*).$

This equivalent form is based on the observation that

$$\begin{bmatrix} F & I \\ I & I \end{bmatrix} = M_1 + M_2 - \mathcal{J},$$

where $M_1$ is the Navier–Stokes term (rightmost matrix in (3.5)) and $M_2$ is the magnetics-velocity term (leftmost matrix in (3.5)). Equivalence follows by recognizing...
that \(\text{SplitPrec}(J, b)\) computes

\[
(M_1^{-1} + M_2^{-1} - M_1^{-1}JM_2^{-1})b \equiv (M_1^{-1}(M_1 + M_2 - J)M_2^{-1})b.
\]

Overall, the algorithm solves the \(2 \times 2\) magnetics-velocity system followed by a correction corresponding to a Navier–Stokes solve with the residual from the previous step as the right-hand side. A practical algorithm is obtained by replacing matrix inverses in \(\text{SplitPrec}(J, b)\) by approximations. Before discussing these, we note that it might be natural to consider reversing the roles of \(M_1^{-1}\) and \(M_2^{-1}\). That is, solve the Navier–Stokes system and then correct with a magnetics-velocity solve. However, we prefer the ordering in \(\text{SplitPrec}(J, b)\) as the velocity component of \(\hat{x}\) satisfies the elliptic incompressibility constraint if the second block component of \(b\) is identically zero.

### 3.2.1. Navier–Stokes equations.

A wide number of Navier–Stokes solvers can be considered to approximate \(M_1^{-1}\). We briefly review some block preconditioners for the Navier–Stokes equations which have demonstrated rapid convergence and good parallel scalability [19, 16, 17, 8]. Several of these are based on the LU factorization

\[
\begin{bmatrix}
F & B^T \\
B & C
\end{bmatrix} = \begin{bmatrix}
I & F^T \\
BF^{-1} & I
\end{bmatrix} \begin{bmatrix}
F & B^T \\
B^T & S
\end{bmatrix},
\]

where again \(S = C - BF^{-1}B^T\). The application of (3.19) requires two matrix solves associated with \(F\). A few multigrid iterations or a few preconditioned Krylov iterations are used to approximate the action of each \(F^{-1}\). When used as a preconditioner for the Navier–Stokes equations (without MHD) one of these \(F^{-1}\)’s can be completely avoided by instead using only the upper triangular factor as a preconditioner [43].

An effective preconditioner requires a relatively inexpensive approximation to \(S^{-1}\) (denoted here by \(\tilde{S}^{-1}\)) yielding the nonreduced version

\[
\begin{bmatrix}
F & B^T \\
B & C
\end{bmatrix} \approx \begin{bmatrix}
I & F^T \\
BF^{-1} & I
\end{bmatrix} \begin{bmatrix}
F & B^T \\
B^T & \tilde{S}
\end{bmatrix},
\]

where [18, 17, 8] discuss several possibilities for \(\tilde{S}^{-1}\). Here, we employ the pressure convection-diffusion (PCD) or SIMPLEC approximations which also require a few multigrid iterations to approximate an embedded Poisson-like solve. The Dirichlet boundary conditions for the pressure convection and diffusion operators used in PCD are chosen to match the Dirichlet boundary conditions set on the pressure variable. Note that this is different from what is done in [18] and [20], and could certainly have impact on the preconditioner. However, to keep this study bounded we have chosen not to explore this issue further.

### 3.2.2. Magnetics-velocity solve.

An approximate inverse of the magnetics-velocity \(2 \times 2\) operator required by (3.5) can also be defined with the help of a block LU factorization

\[
\begin{bmatrix}
F & Z \\
Y & D
\end{bmatrix} = \begin{bmatrix}
I & F \\
YF^{-1} & I
\end{bmatrix} \begin{bmatrix}
F & Z \\
\hat{P}
\end{bmatrix},
\]

where again \(\hat{P} = D - YF^{-1}Z\). To approximate \(\hat{P}^{-1}\), we expand on ideas used in preconditioning the Navier–Stokes equations and also the work of Chacón and coworkers on parabolization of the MHD equations [31, 5].

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
A first \( \hat{P} \) approximation is motivated by the PCD preconditioner for Navier–Stokes. Begin by assuming the following commuting expression holds:

\[
\left( \frac{\partial}{\partial t} + \mathbf{w} \cdot \nabla - \frac{\eta}{\mu_0} \nabla^2 \right) \nabla A_z \cdot \approx \nabla A_z \cdot \left( \frac{\partial}{\partial t} + \mathbf{w} \cdot \nabla - \frac{\mu}{\rho} \nabla^2 \right).
\]

Equation (3.22) is exact if \( \mu/\rho = \eta/\mu_0 \) and \( \nabla A_z \) does not vary spatially, e.g., \( A_z \) is a linear function of the coordinates (a constant \( \mathbf{B} \) field). While generally not true, the exactness of commuting has not been a difficulty in the context of Navier–Stokes [8]. The continuous commuting argument motivates a discrete commuting condition

\[
DQ_a^{-1} Y \approx YQ_u^{-1} F.
\]

The matrices \( Q_u \) and \( Q_a \) are discrete mass operators for the velocity and magnetics spaces whose inverse takes functionals and maps them to functions. They are necessary because an operator discretized by finite elements maps functions to functionals. Substituting the discrete commuting condition into \( \hat{P} \) yields

\[
\hat{P} = D - YF^{-1} Z \approx D - (Q_aD^{-1}YQ_u^{-1})Z = Q_aD^{-1} (DQ_a^{-1}D - YQ_u^{-1}Z)
\]

corresponding to the following inverse approximation:

\[
\hat{P}^{-1}_{Comm} = (DQ_a^{-1}D - YQ_u^{-1}Z)^{-1} DQ_a^{-1}
\]

(here the subscript \( \text{Comm} \) stands for “commuting”). The matrix \( DQ_a^{-1}D - YQ_u^{-1}Z \) is formed explicitly (using lumped mass approximations for \( Q_a \) and \( Q_u \)) and its inverse is approximated by AMG. This approach is attractive because beyond the Jacobian operator it only requires the mass matrix. However, the matrix-matrix multiplication does lead to a matrix with many nonzeros per row. These wide stencil matrices are somewhat nonstandard for AMG and may pose problems. Another potential issue revolves around stabilization where difficulties have been observed with forming effective discrete preconditioners based on operators modified for stabilization with Navier–Stokes [8].

A similar approximation for \( \hat{P} \) can be formulated by starting with the continuous PDE operators and developing Schur complement approximations. To do this, we consider a simplified set of strong form equations representing the magnetics-velocity coupling so as to derive a continuous analog of the \( DQ_a^{-1}D - YQ_u^{-1}Z \) operator. After derivation of this form, a PCD-like argument will be used to motivate a discrete approximation of this continuous operator. The equations

\[
\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} - \frac{\mu}{\rho} \nabla \cdot \nabla \mathbf{u} - \frac{1}{\rho \mu_0} (\nabla \times \mathbf{B}) \times \mathbf{B} = 0,
\]

\[
\frac{\partial \mathbf{B}}{\partial t} - \nabla \times (\mathbf{u} \times \mathbf{B}) + \nabla \times \frac{\eta}{\mu_0} \nabla \times \mathbf{B} = 0,
\]

correspond to the discrete magnetics-velocity system. Using a linear perturbation analysis and applying parabolization ideas [31, 5], we produce a wave propagation system that contains the significant coupling for the magnetics and velocity systems.
For this, the velocity and magnetic fields are written as small perturbations about known spatially invariant and stationary background functions denoted $u_0$ and $B_0$:

\begin{align}
    u &= u_0 + u_1, \\
    B &= B_0 + B_1.
\end{align}

In this analysis, we make the common assumption of a perturbation to a stationary plasma (i.e., $u_0 = 0$) and a constant background magnetic field (i.e., $B_0 = \text{const}$) [23, 3]. Finally the development considers the hyperbolic limit and for this reason the dynamic viscosity and magnetic resistivity are zero (removing the second-order dissipative terms) [23, 3]. Linearizing around $u_0$ and $B_0$ and dropping quadratic and higher-order perturbations gives the linearized set of coupled PDEs:

\begin{align}
    \frac{\partial u_1}{\partial t} - \frac{1}{\rho \mu_0} (\nabla \times B_1) \times B_0 &= 0, \\
    \frac{\partial B_1}{\partial t} - \nabla \times (u_1 \times B_0) &= 0.
\end{align}

The velocity is now eliminated in favor of the magnetic field (this is the continuous analog of a discrete Schur complement). Taking the time derivative of the magnetics equation and substituting the expression for the time derivative of velocity yields

\begin{align}
    \frac{\partial}{\partial t} \left( \frac{\partial B_1}{\partial t} - \nabla \times (u_1 \times B_0) \right) &= \frac{\partial^2 B_1}{\partial t^2} - \frac{1}{\rho \mu_0} \nabla \cdot (\nabla \times B_1) \times B_0 = 0.
\end{align}

The formulation of interest here is the 2D vector potential, so one further step is required to derive the continuous analog of a discrete Schur complement. Taking the time derivative of the magnetics equation and substituting the expression for the time derivative of velocity yields

\begin{align}
    \frac{\partial}{\partial t} \left( \frac{\partial B_1}{\partial t} - \nabla \times (u_1 \times B_0) \right) &= \frac{\partial^2 A_{z1}}{\partial t^2} - \frac{\|B_0\|^2}{\rho \mu_0} \nabla \cdot \nabla A_{z1} = 0.
\end{align}

This expression is an isotropic wave equation propagating with wave speed $\sqrt{\|B_0\|^2/\rho \mu_0}$ which matches the Alfvén wave speed. The result of an isotropic wave equation was the direct consequence of the desire to develop the continuous Schur complement for the reduced magnetics-velocity $2 \times 2$ subsystem, that does not include the enforcement of the incompressibility constraint, in our split preconditioner. However, it needs to be pointed out that this dynamics for the $2 \times 2$ magnetics-velocity subsystem preconditioner does not affect the true anisotropic Alfvén wave physics that is enforced by the convergence of the Newton–Krylov method to the solution of the incompressible resistive MHD equations (1.1)–(1.3).

The continuous isotropic wave equation derived above, corresponds directly to $DQ_a^{-1}D - YQ_u^{-1}Z$. The $DQ_a^{-1}D$ term is a second-order time derivative, and $YQ_u^{-1}Z$ maps to the continuous operator $\frac{\|B_0\|^2}{\rho \mu_0} \nabla \cdot \nabla$. This leads to a second Schur complement approximation

\begin{align}
    \hat{P}_{CSC}^{-1} &= \left( DQ_a^{-1}D - \frac{\|B_0\|^2}{\rho \mu_0} \nabla \cdot \nabla \right)^{-1} DQ_a^{-1},
\end{align}

where the second-order spatial derivative will be discretized via finite elements using piecewise bilinear basis functions (here the subscript CSC stands for “continuous
Schur complement”). The Dirichlet boundary conditions for this operator are the same as those applied to the vector potential equation; the remaining boundaries are treated naturally. Similarly to \( \hat{P}_{\text{Comm}}^{-1} \), the operator \( DQ_u^{-1}D - \frac{[B_{0\rho}]^2}{\rho_{0\rho}} \nabla \cdot \nabla \) is formed explicitly and AMG is used to approximate its inverse. This \( \hat{P} \) approximation is similar to the previous one, but it does avoid the matrix-matrix multiplication associated with \( YQ_u^{-1}Z \) which can lead to a nonstandard matrix for AMG.

A final \( \hat{P} \) approximation is obtained by assuming that the time derivative term dominates the velocity convection diffusion operator. More precisely, the following assumption

\[
F \approx \frac{1}{\Delta t} Q_u
\]

gives rise to

\[
\hat{P}_{\text{Diag}} = D - \Delta t YQ_u^{-1}Z
\]

(here the subscript \( \text{Diag} \) stands for a “diagonal approximation of \( F^{-1} \)”). This approximation is similar to those employed by the SIMPLE family of methods for incompressible Navier–Stokes. \( \hat{P}_{\text{Diag}} \) is formed explicitly using a lumped version of \( Q_u \) and then approximately inverted via AMG.

### 3.3. MHD Neumann series preconditioner.

The previous section focused on avoiding nested Schur complements. Here, we develop an alternative which forms a simplified nested Schur complement approximation giving rise to a preconditioner based directly on (3.3). In particular, let

\[
F_{\text{Neu}} = \text{AbsRowSum}(F),
\]

\[
S_{\text{Neu}} = C - BF_{\text{Neu}}^{-1}B^T,
\]

\[
P_{\text{Neu}} = D - YF_{\text{Neu}}^{-1}\text{AbsRowSum}(I + B_{\text{Neu}})^{-1}BF_{\text{Neu}}^{-1}Z.
\]

The notation \( \text{AbsRowSum}(H) \) denotes a diagonal approximation of an \( n \times n \) matrix \( H \) where the \((i, i)\)th entry of this diagonal matrix is given by \( \sum_{j=1,...,n} |H_{ij}| \). These drastic approximations rely on a Neumann series expansion for each of the inverse operators. The end result is a preconditioner based on

\[
M_{\text{SIMPLEC}} = \begin{bmatrix}
F & B^T & Z \\
S_{\text{Neu}} & -BF_{\text{Neu}}^{-1}Z & P_{\text{Neu}}
\end{bmatrix},
\]

where any algebraic preconditioner can be used to approximate the required inverses of \( F, S_{\text{Neu}}, \) and \( P_{\text{Neu}} \). The lower triangular factor of (3.1) is dropped. Again, this follows Murphy, Golub, and Wathen [43], where it is shown that the lower triangular factor has little effect on convergence.

### 4. Results.

In this initial study, a number of the ABF preconditioners described above are evaluated and compared to standard domain decomposition and an aggressive coarsening AMG preconditioner described in section 4.1. The evaluation considers two different computational MHD problems that are representative of flow-dominated and magnetics-dominated regimes. A study providing a more thorough algorithmic description and an initial comparison of the domain decomposition and aggressive
The first problem presents a hydromagnetic Kelvin–Helmholtz (HMKH) problem used to model the effects of a sheared magnetic field interacting with an unstable velocity shear layer. The parameters of the simulation are chosen so that the magnetics plays a substantive but not dominant role in this problem, so that the magnetic field modifies the vortex dynamics but does not stabilize the shear layer. In the second problem, simulating the coalescence of magnetic islands, the fluid motion is driven by the magnetic field. Here, because the dominant physics is the magnetics, correctly accounting for coupling between the fluid and the magnetic field is critical.

To emphasize the performance of the preconditioners on large-scale problems, we have chosen to do weak scaling studies. In the context of weak scaling, we consider two common approaches for accounting for the scaling between the mesh size and the time step size for fully implicit approaches. For each of these studies, the number of unknowns per processor is fixed and the total number of unknowns grows proportionally to the number of processors used in the problem. The first weak scaling study considers a sequence of increasing total mesh resolutions for which both the flow-velocity and Alfvén-velocity Courant–Friedrichs–Lewy (CFL) numbers are fixed. These parameters are given by

\[
\text{CFL}_V = V_0 \frac{\Delta t}{h}, \quad \text{CFL}_B = V_A \frac{\Delta t}{h},
\]

where \(V_0\) is a characteristic flow velocity and \(V_A = B_0/\sqrt{\rho \mu_0}\) is the Alfvén wave speed. To avoid confusion, we will refer to each of these specifically, or if neither is specified then the ratio \(\Delta t/h\) is used. The second type of weak scaling study uses a fixed time step. In this case, as the mesh is refined the CFL also increases.

The simulations that form this study were run on the Red Sky computer at Sandia National Laboratories. This capacity type machine has 2318 dual socket quad core nodes with a total of 12 GB of RAM per node resulting in 8 cores per node with 1.5 GB of memory per core. Each core is a 2.93 GHz Nehalem processor, and the interconnect is a three-dimensional torus InfiniBand. Because of this, the computational performance to communication speed is not as well balanced above 512 cores as on leadership class capability machines. This has a nonnegligible effect on the CPU time scaling but of course does not affect the algorithmic scaling of the methods, which is the main focus of this initial study (see [36] for a more detailed discussion of these scaling differences).

4.1. Algebraic preconditioners. To serve as a benchmark against which to measure the performance of the block preconditioners, we briefly describe two fully algebraic preconditioners: a common parallel domain decomposition method and a fully coupled AMG preconditioner. The implementation for both of these methods comes from the Trilinos framework [28], and the studies use a GMRES Krylov solver from AztecOO [27]. The preconditioners are primarily from the IFPACK package for domain decomposition and the ML package [57, 21] for AMG.

4.1.1. Schwarz domain decomposition preconditioners. The basic idea of additive Schwarz domain decomposition preconditioners is to decompose the computational domain \(\Omega\) into overlapping subdomains \(\Omega_i\) and then assign each subdomain to a different processor [45]. One application of the algorithm consists of solving on subdomains and then combining these local solutions to construct a global approximation throughout \(\Omega\). The \(i\)th subdomain problem is defined by enforcing homogeneous
Dirichlet boundary conditions on the subdomain boundary, $\partial \Omega_i$. An incomplete factorization, $ILU(k)$, is employed to approximate the solution of the local Dirichlet problems and avoid the large cost of direct factorization [48]. In the minimal overlap case, the algebraic Schwarz method corresponds to block Jacobi, where each block contains all degrees of freedom residing within a given subdomain. Convergence is typically improved by increasing the amount of overlap between subdomains or by increasing the amount of fill in the $ILU(k)$ solver. Both options can lead to increased computational cost per application. For the results presented below, the domain decomposition preconditioners have an overlap of 2 and use $ILU(2)$ as a subdomain solver.

4.1.2. AMG preconditioners. In this paper, only algebraic multigrid methods are considered. These are significantly easier to implement and integrate with complex unstructured mesh simulation software than geometric multigrid methods [57, 51, 53]. Most AMG preconditioners associate a graph with the matrix system being solved. Graph vertices correspond to matrix rows for scalar PDEs, while for PDE systems it is natural to associate one vertex with each nodal block of unknowns (e.g., velocities, pressure, and vector potential at a particular grid point). A graph edge exists between vertex $i$ and $j$ if there is a nonzero in the block matrix which couples $i$’s rows with $j$’s columns or $j$’s rows with $i$’s columns. In some situations, it may be advantageous to omit edges if all entries within the coupling block are small [47].

For this investigation, two different coarsening strategies are used to retain the parallel efficacy through the multigrid hierarchy. Aggressive coarsening (AggC in the figures) is the first of these methods, and is used to precondition the fully coupled MHD system. This preconditioner uses the graph partitioning packages METIS and ParMETIS [30] to subdivide the matrix graph so that each partition is based on a user defined number of graph nodes per aggregate. For the fully coupled system, the nodes per aggregate are chosen so that the graph partitioning algorithm generates somewhat larger aggregates than those typically used. This aggressive coarsening significantly reduces the number of unknowns between consecutive levels, which we find better suited for parallel computations [49, 21]. To compensate for the aggressive coarsening, a somewhat heavyweight Schwarz/$ILU(k)$ smoother (compared to Gauss–Seidel) is used, and, on the coarsest grid, the sparse direct solver KLU [10] is employed.

A second AMG coarsening strategy is employed when developing inverse approximations to the block subsystems of the split preconditioner. This method is a parallel greedy graph aggregation technique which attempts to make an aggregate by taking an unaggregated point and grouping it with all of its neighbors. Thus, it tends to coarsen by a factor of three in each coordinate direction when applied to a standard discretization matrix obtained by a compact stencil on a regular mesh. In addition, a dynamic load-balancing package, Zoltan [12], is used to repartition coarsened operators across processors. This generally improves parallel performance and gives better aggregates on the next coarser level which is obtained by again applying the parallel aggregation technique. A complete discussion of this strategy can be found in [35]. For reference, the AMG parameters used in this study for both the aggressive coarsening and for the blocks systems can be found in Table 4.2.

4.2. Kelvin–Helmholtz instability. The HMKH problem consists of a velocity shear-layer flow that has a sufficient velocity difference to be Kelvin–Helmholtz unstable [22]. This velocity shear layer is subjected to a sheared magnetic field (i.e., a Harris sheet) in the $x$-direction. With no magnetic field, the shear layer is unstable and produces a vortex that rolls up as the simulation progresses. For large enough
For this simulation $Re = 10^3$, $S = 10^4$ and the Alfvén Mach number is 1.5 such that the shear layer is still Kelvin–Helmholtz unstable and the resulting vortex significantly distorts (rolls up) the sheared magnetic field. The two top images are of the initial condition with the sheared magnetic field vectors with color contour background of the magnitude of $B_x$ and the vorticity on the left and right. The remaining images visualize the Kelvin–Helmholtz instability producing a central vortex and bending the magnetic field.

magnetic field strengths, the magnetic stress stabilizes the flow and no vortex is produced. Linear theory indicates that, when the Alfvén Mach number $M_A = V/V_A > 1$, the magnetic field is not strong enough to suppress the KH instability and perturbations on the sheared surface grow. In the example considered here, $M_A > 1$ and the KH instability will distort (roll up) the sheared magnetic field as the vortex forms at the shear surface. Figure 4.1 shows the interaction of a sheared velocity field and magnetic field as a vortex forms.

The HMKH test problem is discretized on the rectangular domain $(x, y) \in [0, 2] \times [0, 1]$ subdivided into uniform quadrilaterals with edge lengths of $1/100^N$, where $N$ is an integer between 1 and 6 (inclusive). Piecewise bilinear basis functions are used for all fields (the Q1 basis). The initial conditions specify a streamwise shear-layer velocity, $V_x$, of $+1.5$ and $-1.5$ in the upper and lower half of the rectangular domain. The
boundary conditions on velocity enforce an inflow velocity of \((+1.5, 0)\) and \((-1.5, 0)\) on the upper half of left \((x = 0)\) and lower half of right \((x = 2)\) boundaries, respectively. On the lower half of left and upper half of right boundaries, respectively, we employ a zero normal-stress boundary condition with \(V_y = 0\). On the upper \((y = 1)\) and lower \((y = 0)\) boundaries there is no penetration \((i.e., V_y = 0)\) and zero tangential stress \((i.e. slip conditions)\). The Alfvén velocity in the upper and lower regions has a magnitude of, \(V_A = 1\). The density \(\rho = 1\), and the reference length scale is taken as the y-direction length \(L = 1\), and the values of the Reynolds number \(Re = \rho V_x L/\mu\) and the Lundquist number \(S = V_A L/\eta\) are given below. The Harris sheet is produced by the vector potential function \((\delta = 0.07957747154595)\)

\[
A_z(x, y, t = 0^+) = \delta \ln \left[ \cosh \left( \frac{y}{\delta} \right) \right] 
\]

(4.2) that is applied on the horizontal boundaries with a natural boundary condition applied at the vertical boundaries of the rectangular domain.

To assess the relative performance of the new block preconditioners against existing algebraic preconditioners as described in section 4.1, we performed weak scaling studies for both fixed CFL and fixed time steps. The simulation was run with a Reynolds number of \(10^3\) and a Lundquist number of \(10^3\). All the results are averaged over 35 time steps starting near \(t = 0.6\ s\), with a required relative tolerance for the nonlinear solver of \(10^{-4}\). The linear solver was iterated until a relative tolerance of \(10^{-3}\) was achieved. In general for these tolerances the average number of Newton steps to achieve convergence, regardless of the preconditioner, ranged between 1 and 2.7 for fixed CFL, and 1 and 2.25 for fixed time step. This problem was run on 1, 4, 16, 64, 256, and 1024 processor cores with uniform grids of size 200 \(\times\) 100, 400 \(\times\) 200, 800 \(\times\) 400, 1600 \(\times\) 800, 3200 \(\times\) 1600, 6400 \(\times\) 3200, respectively. Given that there are four unknowns per mesh node, the total number of unknowns per core is approximately 80,000.

Figures 4.2 and 4.3 show the iteration count and run time as a function of the number of unknowns for a weak scaling study run on the HMKH problem for two values of the CFL \((\Delta t/h = 3.2, 6.4)\). For a CFL proportional to 3.2, the time steps are

\(^{2}\text{It should be noted that larger nonlinear iteration counts of up to 4 occurred on the coarsest mesh resolutions where the solution is underresolved.}\)
Fig. 4.3. Weak parallel scaling with fixed CFL of the average run time of a single Newton solve for the transient HMKH problem with Reynolds number of $10^3$.

Fig. 4.4. Weak parallel scaling with fixed time step of the average iteration count per Newton step of a single Newton solve for the transient HMKH problem with Reynolds number of $10^3$.

[0.032, 0.016, 0.008, 0.004, 0.002, 0.001], and for a CFL proportional to 6.4 the time steps are [0.064, 0.032, 0.016, 0.008, 0.004, 0.002]. Figures 4.4 and 4.5 show the iteration count and run time as a function of the number of unknowns for a weak scaling study run on the HMKH problem for three different fixed time steps ($\Delta t = 0.005, 0.0025, 0.00125$). For the fixed time step scaling study, the fluid CFL increases from left to right with a maximum CFL of 16, 8, and 4 for $\Delta t = 0.005, 0.0025, 0.00125$, \ldots
respectively. For both sets of figures, the legend combined with Tables 4.1 and 4.2 provides the details of each preconditioner. In this case, DD (domain decomposition) and AggC (aggressive coarsening) correspond to the algebraic preconditioners discussed in section 4.1. Crosses on a data point indicate that the simulation made some progress but did not complete for that problem size and processor count. If a data point is missing, then the preconditioner was wholly ineffective and the linear solver failed on the first time step.

The plots in Figure 4.2 show that the domain decomposition method does not scale. The number of linear iterations increases as the processor and unknown count increases. In spite of this, its overall run time performance (see Figure 4.2) seems reasonable; however, it will decay further for large problems run with more unknowns as is clearly evident from the iteration count scaling. The SIMPLEC method does not fare much better, with performance seeming to be highly dependent on mesh resolution for the CFL = 3.2 problem, and total failure for CFL = 6.4. The aggressive coarsening preconditioner (AggC-ILU₆₀) is algorithmically scalable, meaning iteration count is nearly flat with increasing mesh resolution. The run times nearly follow this same trend, with the only blemish occurring for the largest problem size and processor count resulting in a mild run time increase. This is a result of an imbalance in the processor to communication speed on this capacity-type machine. The BlkUp preconditioner performs well for the largest problem sizes (and smallest time steps in this fixed CFL study) with the number iterations decaying with the time step.

Fig. 4.5. Weak parallel scaling with fixed time step of the average run time of a single Newton solve for the transient HMKH problem with Reynolds number of 10³.
A PRECONDITIONER FOR 2D INCOMPRESSIBLE MHD

Table 4.1
ML parameters for the different block preconditioners used. The parameters for the AMG solver for each solve type, velocity, pressure, or magnetics, are specified in Table 4.2.

<table>
<thead>
<tr>
<th>Label</th>
<th>Preconditioner</th>
<th>Schur complement</th>
<th>AMG solver</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIMPLEC</td>
<td>Eq. (3.38)</td>
<td>N/A</td>
<td>NSA-ILU</td>
</tr>
<tr>
<td>BlkUp</td>
<td>Eq. (3.4)</td>
<td>PCD</td>
<td>NSA-ILU</td>
</tr>
<tr>
<td>Split Comm</td>
<td>SplitPrec-NS</td>
<td>PCD $\hat{P}_{\text{Comm}}$ (see (3.25))</td>
<td>NSA-ILU</td>
</tr>
<tr>
<td>Split CSC</td>
<td>SplitPrec-NS</td>
<td>PCD $\hat{P}_{\text{CSC}}$ (see (3.32))</td>
<td>NSA-ILU</td>
</tr>
<tr>
<td>Split Diag</td>
<td>SplitPrec-NS</td>
<td>PCD $\hat{P}_{\text{Diag}}$ (see (3.34))</td>
<td>NSA-ILU</td>
</tr>
</tbody>
</table>

Table 4.2
ML parameters for the different AMG solvers used. All AMG solvers use only one application of the V- or W-cycle. The acronyms NSA and SA stand for nonsmooth aggregation and smooth aggregation, respectively, and "ov" stands for “overlap.”

<table>
<thead>
<tr>
<th>Label</th>
<th>Type</th>
<th>Cycle-type</th>
<th>Smoother</th>
<th>Coarse size nodes/Agg levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>AggC-ILU</td>
<td>NSA</td>
<td>MGW</td>
<td>ILU(2,2)$_2$</td>
<td>500 $N$ 3</td>
</tr>
<tr>
<td>SA-GS</td>
<td>SA(ω = 4/3)</td>
<td>MGV</td>
<td>GS$_2$</td>
<td>25000 – –</td>
</tr>
<tr>
<td>NSA-ILU$_{ab}$</td>
<td>NSA</td>
<td>MGV</td>
<td>ILU(fill=a,ov=b)</td>
<td>25000 – –</td>
</tr>
</tbody>
</table>

size. However, for the range of the smaller problems with a bigger time step, the number of required linear iterations is large, with the method even exhibiting failure for the smallest problem sizes. Finally, the three operator split block preconditioners (Split-Comm, Split-CSC, and Split-Diag) all perform very well exhibiting a decay in the number of linear iterations required per time step. This is not reflected in the run time plots, however. We believe this is a result of the relatively high cost of construction versus application of these preconditioners. Furthermore, the Split-B method takes on the order of 6–8 s longer per nonlinear step, versus Split-Comm and Split-Diag. This is because of the need to construct auxiliary operators in forming the Schur complement. Note, however, that our code is not optimized when building these auxiliary operators and we expect from timing studies that 5 s of the time can be made up with more efficient construction.

For the fixed time step studies (shown in Figures 4.4 and 4.5), the domain decomposition preconditioner again scales poorly. The SIMPLEC method also does poorly as the problem size increases. We attribute this to the poor performance of the Neumann-series-based Schur complement approximations for fluid flow with large CFLs (see [8]). The aggressive coarsening preconditioner, despite performing well on smaller problems, performs poorly for the largest time step and problem sizes. Similarly the Split-Comm block preconditioner does poorly for the largest problem sizes (with the exception of the smallest time step). Further investigation into the reason for this failure for the Split-Comm preconditioner showed that the Schur complement approximation was the issue. Essentially, the AMG preconditioner used to approximately invert the $DQ_a^{-1}D - YQ_u^{-1}Z$ operator was diverging. The remaining block preconditioners (BlkUp, Split-CSC, and Split-Diag) all scale well with only a modest increase in the iteration count over the range of unknown and processor counts. For all time step sizes, all methods have a slight increase in the number of linear iterations with the number of unknowns, with the BlkUp preconditioner suffering the greatest increase. This is consistent with the spectral analysis presented earlier. In
Fig. 4.6. These cropped images contain contour plots for the island coalescence driven magnetic reconnection at times $t = 0.0$ on the left and $10.0$ on the right. The images show isolines of the magnetic potential $A_z$ and filled contours of the current $J_z$. The formation of the x-point and the thin current sheet is clearly evident.

In general, we see that BlkUp, Split-Comm, Split-CSC, and Split-C are all essentially scalable (with nearly flat linear iterations and run time), with the one exception of a robustness issue for Split-Comm on the largest problem size. The scalability of the splitting-type preconditioners indicates that the split preconditioning methodology is a reasonable approach in this case, despite the variation in robustness over different Schur complement approximations. While the scalability of the BlkUp preconditioner appears promising, we will see that, in other problem regimes, the seemingly modest sensitivity with respect to $\Delta t$ is much more significant.

### 4.3. Island coalescence

The island coalescence problem consists of a perturbed Harris sheet magnetic field configuration [2, 32, 50] that introduces two magnetic islands in the plasma as initial conditions. The structure of this perturbation can be seen in the initial condition plot at time $t = 0$ of Figure 4.6 (left) with isolines of $A_z$. The centers of the two islands are referred to as o-points, and the point between them where a thin current sheet is formed is referred to as the x-point. The combined magnetic field produced by the two magnetic islands produces Lorentz forces that pull the islands together. For larger resistivities, the x- and o-points steadily approach each other. For low resistivities, magnetic pressure builds up as the islands approach, and a sloshing or bouncing of the o-point position is encountered that leads to lower reconnection rates (for more details on the physics, see [2, 44]). The right image in Figure 4.6 shows an isoline plot of $A_z$ and filled color contours of the plasma current $J_z$ during the reconnection event at time $t = 10.0$. Clearly evident is the formation of the x-point between the islands, the development of a thin current sheet at the x-point location. The movement of the center of the islands (o-points) towards the x-point is apparent [2, 32].

The initial conditions, $A_0^z$, and the resulting balancing plasma fluid pressure, $P_0$, are given by

\begin{align}
A_0^z(x, y, 0) &= \delta \ln \left[ \cosh \left( \frac{y}{\delta} \right) + \epsilon \cos \left( \frac{x}{\delta} \right) \right], \\
P_0^0(x, y, 0) &= P_0 + \frac{[1 - \epsilon^2]}{2 \left[ \cosh \left( \frac{y}{\delta} \right) + \epsilon \cos \left( \frac{x}{\delta} \right) \right]^2},
\end{align}

where $\delta = 1/(2\pi)$ and $P_0 = 1.0$ (see [50] for more details).

For the island coalescence problem, we exploit its physical symmetry and simulate only the right half of the domain (note that the images in Figure 4.6 are of the full domain and have been cropped to emphasize the magnetic islands and the current sheet). We discretize the rectangular domain, defined on $(x, y) \in [0, 1] \times [-1, 1]$, using uniform quadrilaterals with edge lengths of $2^{-N}$, where $N$ is an integer between 6 and 11 (inclusive). Piecewise bilinear basis functions are used for all fields (the Q1 basis). The boundary conditions are $u \cdot n = 0$ on all boundaries with zero stress in...
the tangential component on each surface. Both the vector potential and pressure Dirichlet condition are set on the $y = \pm 1$ boundaries using the definitions of $A_0^z$ and $P^0$, respectively. For this problem we set the viscosity and resistivity to $10^{-4}$, and set density and magnetic permeability to unity. The characteristic length scale is taken as the distance between o-points, which is the x-dimension length of $L = 1.0$ for the computation of the Lundquist number.

To assess the relative performance of the new block preconditioners against existing algebraic preconditioners as described in section 4.1, we performed weak scaling studies for both fixed time step and fixed CFL simulations. As with the HMKH problems, the details of each preconditioner in the legend can be found in Tables 4.1 and 4.2. Crosses on a data point indicate that the simulation did not complete (i.e., failure of the linear solver caused at least one failure of the nonlinear solver and time step) for that problem size and processor count. All island coalescence results are averaged over 45 time steps starting around $t = 5.75 \text{ s}$ with a required relative tolerance for the nonlinear solver of $10^{-4}$. The linear solver used a relative tolerance of $10^{-3}$. For these tolerances, the average number of Newton steps to achieve convergence regardless of preconditioner ranged between 1.5 and 2.25. This problem was run on 1, 4, 16, 64, 256, and 1024 processor cores with uniform grids of size $64 \times 128$, $128 \times 256$, $256 \times 512$, $512 \times 1024$, $1024 \times 2048$, and $2048 \times 4096$, respectively. Given that there are four unknowns per node, the number of unknowns per core is approximately 33,000.

The set of plots in Figures 4.7 and 4.8 for average linear iterations and time per Newton step show the scaling of the preconditioners with a fixed CFL. In this
case we have an Alfvén wave speed of \( V_A = B_0 / \sqrt{\mu_0 \rho} = 1 \) and therefore \( CFL_A = \Delta t / h = 3.2, 6.4, 12.8 \). From these plots, we see that the BlkUp and the three split block preconditioners exhibit a decay in the total number of linear iterations as a function of the number of unknowns. This is consistent with the results shown for the HMKH problem. Again, as expected, the aggressive coarsening AMG preconditioner maintains flat iteration counts regardless of CFL.

To demonstrate the algorithmic scalability of the various preconditioners for a fixed time step, weak scaling study of the island coalescence problem, we plot the average linear iterations required as a function of problem size in Figure 4.9. In this weak scaling study the time steps that were considered were \( \Delta t = 0.00625, 0.0125, 0.025, 0.05 \). This sequence results in an increasing CFL with increased mesh resolution (smaller \( \Delta x \)) and a maximum \( CFL_A = 102.4 \). From the sequence of plots, we see that only the aggressive coarsening algorithm iteration count is relatively independent of the time step parameter. Further, the aggressive coarsening preconditioner is essentially scalable for all time steps. Bear in mind however, this algorithm is heavily reliant on the equal-order approximation used in the stabilized simulation, and is not applicable to mixed discretizations. The domain decomposition and SIMPLEC algorithms do not scale well, even for the smallest time step. Furthermore, the applicability of the Neumann series approximation used in SIMPLEC is strained with increased CFL regardless of time step. Compared to the split methods, the average number of linear iterations required for the BlkUp block preconditioner is highly dependent on the time
Fig. 4.9. Weak parallel scaling of the average number of linear iterations for the transient island coalescence problem with Lundquist number of $10^4$. The time steps are fixed for each plot.

step, with the BlkUp preconditioner failing to converge for the largest problem sizes using $\Delta t = 0.05$. In contrast, the split methods have reasonably good algorithmic scalability, with all three approximations scaling well over all problem sizes. However, here the importance of accurate approximation and solving the magnetics-velocity Schur complement is clear. There appears to be roughly a 1.5–2 times variation in the linear iteration counts between these methods for the largest time steps, a relatively modest increase considering the problem sizes have increased by 1024 times. The most consistent of these for this problem, Split-Comm, uses an algebraic commuting argument. However, as we saw in the HMKH problem, Split-Comm can suffer from lack of robustness.

The run time per linear iteration corresponding to the weak scaling study described above can be found in Figure 4.10. Here again, as with the HMKH problem, there is a strong positive correlation between algorithmic scalability and the scalability of the run time for each method. One thing of note is that despite a roughly 3–4 times increase in the linear iteration count for increasing time step, the run times of the block preconditioners, especially Split-Comm, remains roughly twice that of the aggressive coarsening method. Furthermore, all algorithmically scalable methods suffer a modest increase in run time with weak scaling. This is due to the imbalance of communication and computation on this capacity machine. Again, we point out that our block code is not fully optimized for the block preconditioners, and there is likely further time savings for these methods that could be achieved.
5. Conclusions. We have presented three new ABF preconditioners for a stabilized vector potential incompressible resistive MHD formulation. Over the problems considered, the most effective of these is based on an operator-split approximation which decomposes the $3 \times 3$ linear operator into two $2 \times 2$ operators. One of the two operators represents the fluid components only coupling the velocity and pressure fields. The inverse of this operator can be easily approximated with existing block preconditioners for Navier–Stokes. The second operator explicitly couples the magnetics and velocity systems. For this operator, we proposed several approximations to its inverse using block factorizations. These differ primarily in how the Schur complement is approximated. A block upper triangular preconditioner that ignores the effects of the velocity-magnetics coupling was also developed. The final block preconditioner is based on a full block LU factorization of the Jacobian operator utilizing relatively simple Neumann series expressions (with robustness concerns) to construct approximations of the nested Schur complements.

The parallel performance of the ABF preconditioners was compared to a domain decomposition method and a fully coupled AMG preconditioner for two challenging transient problems. These weak scaling studies demonstrate that the operator-split preconditioner is algorithmically scalable for both fixed CFL and fixed time step studies. Throughout these studies, the run time performance, while not optimized, is within 2–3 times of the fully coupled multigrid algorithm. We believe that the advantage of the operator-split method is that it is procedurally applicable to any discretization that shares the same structure as the one considered here.
tably this includes mixed discretizations where the fully coupled multigrid solver is not appropriate. Exploration of the performance of these preconditioners to various mixed discretizations is the subject of future work. The block upper triangular block factorization method also proved to be effective for some parameter choices (i.e., weakly coupled momentum and magnetics). However, for other less favorable parameter choices (i.e., strongly coupled momentum and magnetics), this method shows a severe dependence on time step size which results in poor performance. The block factorization method that uses drastic Neumann series approximations for the Schur complement also suffers from a strong time step dependence. It suffers serious performance degradation for increasing CFL numbers, though for large scale problems with small CFL numbers this preconditioner can be effective.

6. Appendix.

6.1. Stabilization parameters. The specific definitions of the stability parameters are provided in Table 6.1 for momentum and the vector potential. The specific form of the $\tau$’s are an adaptation of the quadratic form by Shakib [54] for the Navier–Stokes equations and the Lorentz form stabilization operator from Codina and Hernandez-Silva [7] for a resistive MHD system. Additional forms for the stability parameters are discussed in [13] and the references contained therein for Navier–Stokes. The contribution to the $\hat{\tau}_m$ operator for the Lorentz force term uses the same form as in [7] with $C_2 = 10$. It should be pointed out that the multidimensional effect of convection is incorporated into the stability parameters by the use of the contravariant metric tensor, $G_c$ (6.1), of the transformation from local element coordinates $\{\zeta_{\alpha}\}$ to physical coordinates $\{x_i\}$:

$$[G_c]_{ij} = \frac{\partial \zeta_{\alpha}}{\partial x_i} \frac{\partial \zeta_{\alpha}}{\partial x_j}. $$

Shakib [54] considers the one-dimensional limiting case of this multidimensional definition for the advection-diffusion equation and presents a comparison with the original SUPG technique. A more detailed description of the resistive MHD stabilized formulation and the stability parameters used in our initial study can be found in [50].

<table>
<thead>
<tr>
<th>Table 6.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition of stabilization parameters used in stabilized equations, which use the contravariant metric tensor $G_c$ (6.1) to define an appropriate element-level length scale. In this study $C_1 = 1$ and $C_2 = 10$.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Momentum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\tau}_m = \left[ \left( \frac{2\mu}{\Delta t} \right)^2 + \rho^2 v G_c \cdot v + C_1^2 \mu^2 | G_c |^2 + C_2^2 \frac{\rho}{\mu_0} | B |^2 | G_c | \right]^{-\frac{1}{2}} $</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Z-component vector potential</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\tau}_{A_z} = \left[ \left( \frac{2}{\Delta t} \right)^2 + v G_c \cdot v + C_1^2 \eta^2 | G_c |^2 \right]^{-\frac{1}{2}} $</td>
</tr>
</tbody>
</table>

3It should be noted that in [50] there were typographical errors that incorrectly represented the element length scale contributions for the diffusion and Lorentz force term and also the order of contribution of $\| B \|$ in the Lorentz force term. The definition above has the correct scaling.
REFERENCES


A PRECONDITIONER FOR 2D INCOMPRESSIBLE MHD


