Magnetic Vector Potential Formulation for Variational Data Assimilation in Ideal Magnetohydrodynamics

J. H. Arnal* and C. P. T. Groth
Corresponding author: `jose.arnal@mail.utoronto.ca`
University of Toronto Institute for Aerospace Studies, Canada

Abstract: Magnetohydrodynamics (MHD) is a well established description of electrically conducting fluids. A central challenge in computational MHD is the enforcement of the divergence-free condition for the magnetic field which must be satisfied to avoid unphysical behavior and loss of numerical stability. This challenge persists in the setting of variational data assimilation (DA), the process of systematically combining observational data with a physical model to produce an improved estimate of the system state. In this context, magnetic divergence errors arise not only from numerical discretization but also from the assimilation of sparse and noisy observational data. The present study proposes a novel formulation of ideal MHD-based variational DA that eliminates the introduction of such errors by construction. The central idea is to infer an underlying magnetic vector potential, rather than the magnetic field directly, thereby guaranteeing satisfaction of the divergence-free condition. The approach is formulated as a control-variable transform within a standard upwind finite-volume discretization and adjoint-based optimization framework, and requires only minimal modification of existing variational DA frameworks. The method is assessed through observing system simulation experiments based on two highly non-linear plasma flows. Relative to a standard baseline variational DA formulation, the proposed approach achieves comparable accuracy for dense observations and markedly improved accuracy for sparse observations, while introducing no divergence errors through assimilation.

Keywords: Magnetohydrodynamics, Data Assimilation, Adjoint Method, Finite-Volume Method

1. Introduction and Motivation

Magnetohydrodynamics (MHD) is a well established phenomenological description of electrically conducting fluids, with applications in space physics [1], geophysics [2], and nuclear fusion [3]. A central challenge in computational MHD is the enforcement of the divergence-free condition for the magnetic field, $\nabla \cdot \vec{B} = 0$. Although not an explicit transport equation of the MHD system, it appears as an involution constraint that is always satisfied given compatible initial and boundary conditions. At the discrete level, however, the solenoidal property can be severely violated if discretization schemes are not constructed with special care. Large non-zero magnetic divergence errors can result in unphysical behavior and may compromise numerical stability. As a consequence, a variety of strategies have been proposed and developed to control divergence errors in computational MHD, including Powell's formulation [4, 5], projection methods [6], and constrained transport schemes [7].

Recently, considerable attention has been directed toward the application of variational data assimilation (DA) to plasma flows governed by ideal MHD descriptions. In particular, the authors were the first to apply variational DA to ideal MHD [8, 9]. Data assimilation is the process of systematically combining observational data with a physical model to produce an improved estimate of the system state. In the variational approach, this is formulated as a partial-differential-equation-constrained optimization problem, in which a cost functional measuring the misfit between model predictions and observations is minimized subject to the governing equations acting as constraints. In the context of solar wind

*Currently Postdoctoral Research Fellow, University of Cambridge Department of Engineering, UK forecasting, for example, MHD-based variational DA can yield significant predictive improvements compared to that of the forward modelling alone [10, 9].

Analogous to the challenges encountered in forward MHD modelling, the enforcement of $\nabla\cdot\vec{B} = 0$ emerges as a critical problem in the DA setting as well. The distinction, however, is that while forward modelling techniques must treat divergence errors introduced by numerical discretization, DA must additionally handle errors that arise from the data because of sparse spatial coverage, noisy measurements, and issues inherent to the DA algorithm. The present study proposes a novel formulation of ideal MHD-based variational DA that eliminates the introduction of magnetic divergence errors by construction. The central idea is to infer an underlying magnetic vector potential, rather than the magnetic field directly, thereby guaranteeing satisfaction of the divergence-free condition.

The organization of the remainder of this article is as follows. Section 2 presents the ideal MHD description of plasmas and the upwind finite-volume scheme used here for obtaining numerical approximations of unsteady plasma flows. Section 3 first describes the Bayesian inversion approach typically employed in variational DA and its application to compressible plasmas. This is followed by the formulation of a control-variable transform, based on the magnetic vector potential representation of magnetic fields, to enforce the $\nabla\cdot\vec{B} = 0$ condition during variational DA. Several numerical experiments involving the assimilation of synthetic plasma measurements with and without the novel vector potential formulation are then presented and discussed in Section 4. Finally, Section 5 draws conclusions, provides a summary, and suggests recommendations for future research.

2. Forward Modelling of Compressible Plasmas

2.1 Governing Equations

The ideal MHD equations describe a perfectly electrically conducting plasma where the mixture of ion and electron species is treated as a single fluid that is fully-ionized, inviscid, quasi-neutral, and possesses a Maxwellian particle distribution described by a single temperature. Important processes that influence thermodynamic behavior–such as kinetic effects, resistivity, viscosity, thermal conduction, and multi-species interaction–are thus neglected, while low-frequency, large-scale phenomena is presumed. Despite these assumptions, the ideal MHD description has been widely adopted for the computational modelling of plasmas.

In weak conservation form, the ideal MHD equations are given by

$$\frac{\partial \boldsymbol{U}}{\partial t} + \nabla\cdot\boldsymbol{F} = \boldsymbol{S}\,, \tag{1}$$

where $\boldsymbol{U}$ is the solution vector of conserved quantities, $\boldsymbol{F}$ is the system flux dyad, and $\boldsymbol{S}$ is a vector of source terms, which are given by

$$\boldsymbol{U} = \begin{bmatrix} \rho \\ \rho\vec{u} \\ \vec{B} \\ e \end{bmatrix}, \quad \boldsymbol{F} = \begin{bmatrix} \rho\vec{u} \\ \rho\vec{u}\vec{u} - \vec{B}\vec{B} + (p + \frac{\vec{B}\cdot\vec{B}}{2})\vec{\vec{I}} \\ \vec{B}\vec{u} - \vec{u}\vec{B} \\ \vec{u}(e + p + \frac{\vec{B}\cdot\vec{B}}{2}) - \vec{B}(\vec{u}\cdot\vec{B}) \end{bmatrix}, \quad \boldsymbol{S} = -\nabla\cdot\vec{B}\begin{bmatrix} 0 \\ \vec{B} \\ \vec{u} \\ \vec{u}\cdot\vec{B} \end{bmatrix}. \tag{2}$$

In these definitions, ρ is the plasma mass density, $\vec{u} = [u, v, w]^\intercal$ is the velocity vector, p is the thermal pressure, $\vec{B} = [B_x, B_y, B_z]^\intercal$ is the magnetic field, e is the total plasma energy, and $\vec{\vec{I}}$ is the identity dyad. The source term, $\boldsymbol{S}$, is proportional to $\nabla\cdot\vec{B}$ and therefore vanishes in the continuous setting; its role is discussed in Section 2.3. Magnetic field units are chosen here such that the magnetic permeability of vacuum, μ_0, is taken to be equal to unity. Moreover, non-dimensional units are adopted in all calculations considered here.

A polytropic (thermally and calorically perfect gas) approximation is assumed for describing the thermodynamic behavior of the plasma and thus the total energy and plasma pressure are related according to

$$e = \frac{p}{\gamma - 1} + \frac{\rho|\vec{u}|^2}{2} + \frac{|\vec{B}|^2}{2}\,, \tag{3}$$

where γ is the corresponding assumed constant ratio of specific heat capacities. The ideal gas equation of state, $T = \gamma p/\rho$, is also assumed, where T is the non-dimensional plasma temperature.

2.2 Upwind Finite-Volume Discretization

A Godunov-type, second-order, upwind, shock capturing, finite-volume scheme [11] is considered here for the approximate solution of the ideal MHD equations. Finite-volume methods of this type have proven to be very effective for the solution of hyperbolic partial differential equations, including the ideal MHD equations. This spatial discretization scheme is employed here with grids composed entirely of hexahedral elements yielding a semi-discrete form given by

$$\frac{\mathrm{d}\boldsymbol{U}_K}{\mathrm{d}t} = -\frac{1}{\mathcal{V}_K}\sum_{q=1}^{N_q}(\boldsymbol{F}_q \cdot \hat{n}_q \mathcal{A}_q)_K + \boldsymbol{S}_K \equiv \boldsymbol{R}_K\,, \tag{4}$$

where the sum is over the $N_q = 6$ quadrature points located on the surface mid-points of the computational cell, K. Here, $\boldsymbol{U}_K$ denotes the cell-averaged solution vector, $\boldsymbol{F}_q$ is the upwind value of the solution fluxes evaluated at the qth quadrature point along the boundary of the cell, $\boldsymbol{S}_K$ denotes the cell-averaged source vector, $\mathcal{V}_K$ is the cell volume, $\mathcal{A}_q$ is the hexahedral face area associated with each quadrature point, and $\hat{n}_q$ denotes the outward pointing unit normal vector to the hexahedral cell face, ∂K_q. The right-hand side of Eq. (4) defines the steady-state discrete residual, denoted by $\boldsymbol{R}_K$, which encapsulates the spatial discretization of the governing equations and therefore vanishes in steady state conditions.

The upwind fluxes are evaluated here via the approximate solution of Riemann problems defined in terms of the primitive solution quantities, $\boldsymbol{W} = [\rho, \vec{u}, \vec{B}, p]^\intercal$, evaluated at each side of a computational cell interface via the Harten-Lax-van Leer-Einfeldt (HLLE) approximate Riemann solver [12]. Second-order spatial accuracy is achieved by piece-wise limited linear reconstruction of the primitive solution vector. The least-squares method [13] is employed for the evaluation of primitive solution gradients together with the Venkatakrishnan slope limiter [14] for the avoidance of spurious oscillations in the vicinity of solution discontinuities.

The cell-averaged source term, $\boldsymbol{S}_K$, is approximated by

$$\boldsymbol{S}_K = \frac{\widetilde{\boldsymbol{S}}_K}{\mathcal{V}_K}\sum_{q=1}^{N_q}(\vec{B}_q \cdot \hat{n}_q \mathcal{A}_q)_K\,, \quad \text{with} \quad \widetilde{\boldsymbol{S}}_K = -\begin{bmatrix} 0 \\ \vec{B} \\ \vec{u} \\ \vec{u}\cdot\vec{B} \end{bmatrix}_K\,, \tag{5}$$

where the face magnetic field, $\vec{B}_q$, is obtained by averaging the piecewise-linear magnetic-field reconstructions evaluated on either side of the cell interface.

For the unsteady initial-boundary-value problems of interest here, the semi-discrete form of the ideal MHD equations (Eq. (4)) is solved by adopting a method-of-lines approach and applying a strong stability preserving (SSP), second-order, Runge-Kutta, time-marching scheme [15, 16] given by

$$\widetilde{\boldsymbol{U}}^{n+1} = \boldsymbol{U}^n + \Delta t \boldsymbol{R}(\boldsymbol{U}^n)\,, \tag{6}$$

$$\boldsymbol{U}^{n+1} = \frac{1}{2}(\boldsymbol{U}^n + \widetilde{\boldsymbol{U}}^{n+1} + \Delta t \boldsymbol{R}(\widetilde{\boldsymbol{U}}^{n+1}))\,, \tag{7}$$

where $\boldsymbol{U}^n$ stores the cell-averaged conserved solution vectors for the entire grid at the n^{th} time step.

The SSP time-marching scheme ensures monotone solutions for an appropriately restrictive choice of the global time step, Δt, satisfying the usual Courant-Friedrichs-Lewy (CFL) criteria [17].

As with the steady residual, $\boldsymbol{R}$, we define the unsteady residual, $\widetilde{\boldsymbol{R}}$, by rearranging Eqs. (6) and (7) to yield

$$\widetilde{\boldsymbol{R}} \equiv \frac{\boldsymbol{U}^{n+1} - \boldsymbol{U}^{n}}{\Delta t} - \frac{\boldsymbol{R}(\boldsymbol{U}^{n}) + \boldsymbol{R}(\widetilde{\boldsymbol{U}}^{n+1})}{2} = \boldsymbol{0}\,. \tag{8}$$

The definition of the unsteady residual encodes the fully discrete ideal MHD equations and will be used as a constraint in the DA method to be described in Section 3.

2.3 Magnetic Solenoidality

The source term, $\boldsymbol{S}$, is typically neglected in the continuous ideal MHD equations, as it vanishes due to Gauss's law for magnetism,

$$\nabla \cdot \vec{B} = 0\,. \tag{9}$$

While Eq. (9) is not explicitly part of the ideal MHD system, the magnetic induction equation satisfies an involution constraint, $\partial(\nabla \cdot \vec{B})/\partial t = 0$, which ensures the magnetic field remains divergence-free for all time provided the initial data are divergence-free.

Nevertheless, the source term, $\boldsymbol{S}$, must be retained to preserve the symmetrizable form of the ideal MHD system [18]; this form admits an entropy conservation law and is Galilean invariant [19, 20]. Moreover, at the discrete level, the involution constraint is not satisfied under general discretization strategies. By retaining the source term, $\boldsymbol{S}$, Powell [4] showed that divergence errors are advected along streamlines out of the computational domain, which, in many cases, prevents the accumulation of large $\nabla \cdot \vec{B}$ errors and consequently improves numerical stability.

3. Variational Data Assimilation for Ideal Magnetohydrodynamics

3.1 Standard Variational Formulation

We consider initial-boundary-value problems where the initial data contains stochastic errors which degrade the prediction of unsteady plasma flows. A variational DA approach is adopted to assimilate *in situ*, noisy, and sparse plasma measurements and thereby reconstruct the true plasma flow field evolution. We limit our analysis to Gaussian error statistics where the available initial data, $\check{\boldsymbol{U}}^0$, and observational measurements, $\boldsymbol{d}^n$, take the form

$$\check{\boldsymbol{U}}^0 = \boldsymbol{U}_{\mathrm{t}}^0 + \boldsymbol{\xi}\,, \qquad \text{with} \quad \boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{U}})\,, \tag{10}$$

$$\boldsymbol{d}^n = \boldsymbol{\mathcal{H}}(\boldsymbol{U}_{\mathrm{t}}^n) + \boldsymbol{\epsilon}^n\,, \qquad \text{with} \quad \boldsymbol{\epsilon}^n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{d}})\,, \tag{11}$$

where $\boldsymbol{U}_{\mathrm{t}}^n$ denotes the true plasma state at the n^{th} time-step, and $\boldsymbol{\mathcal{H}}$ is the observation operator mapping the state onto measurement space. Here, $\boldsymbol{\xi}$ and $\boldsymbol{\epsilon}^n$ denote Gaussian errors in the initial data and observational measurements, respectively, with known covariance matrices, $\boldsymbol{\Sigma}_{\boldsymbol{U}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{d}}$, respectively.

Adopting a Bayesian view of probability, the variational DA approach seeks the maximum *a posteriori* (MAP) estimate of the initial data, $\boldsymbol{U}^0$, given the measurements, $\boldsymbol{d}^n$. Under the assumption of Gaussian error statistics, the MAP estimate is given by the solution of the following optimization problem

$$\begin{aligned} \underset{\boldsymbol{U}^0}{\arg\min} \quad & J = \frac{1}{2}\sum_{n=1}^{N} ||\boldsymbol{\mathcal{H}}(\boldsymbol{U}^n) - \boldsymbol{d}^n||^2_{\boldsymbol{\Sigma}_{\boldsymbol{d}}^{-1}} + \frac{1}{2}||\boldsymbol{U}^0 - \check{\boldsymbol{U}}^0||^2_{\boldsymbol{\Sigma}_{\boldsymbol{U}}^{-1}}\,, \\ \text{s.t.} \quad & \tilde{\boldsymbol{R}} = \boldsymbol{0}\,. \end{aligned} \tag{12}$$

The first term of the cost function, J, measures the distance between modeled and measured plasma properties, weighted by the precision of observations, while the second measures the squared difference between the posterior and prior initial plasma state, weighted by the precision of the model, and may be regarded as a regularization term that naturally arises from the Bayesian formulation. A solution to Eq. (12) therefore optimally balances the information contained in the observational data and the *a priori* estimate, while respecting the physical description of plasmas encoded in the ideal MHD equations.

The optimization problem formulated in Eq. (12) is not typically solved directly as stated. Instead, a control-variable transformation (CVT) is generally employed to precondition the optimization problem and thereby achieve faster numerical convergence [21, 22]. With this approach, the initial data, $\boldsymbol{U}^0$, is written in terms of a new control variable, $\boldsymbol{V}$, according to

$$\boldsymbol{L}\boldsymbol{V} = (\boldsymbol{U}^0 - \check{\boldsymbol{U}}^0)\,, \tag{13}$$

where $\boldsymbol{L}$ is the lower Cholesky factor of the background covariance matrix, $\boldsymbol{\Sigma}_{\boldsymbol{U}}$, which is defined by

$$\boldsymbol{L}\boldsymbol{L}^\intercal = \boldsymbol{\Sigma}_{\boldsymbol{U}}\,. \tag{14}$$

Under the CVT, the DA optimization problem is solved in terms of $\boldsymbol{V}$ as opposed to $\boldsymbol{U}^0$. Numerical solutions of this optimization problem are obtained with gradient-based optimization and adjoint methods as is detailed further in Section 3.3.

3.2 Magnetic Vector-Potential Formulation

A significant limitation of the preceding, standard, variational DA formulation is the lack of enforcement of the $\nabla \cdot \vec{B} = 0$ condition. As outlined in Section 2.3, the ideal MHD equations only enforce $\nabla \cdot \vec{B} = 0$ through involution, meaning that the time evolution of a non-solenoidal initial magnetic field, $\vec{B}^0$, is fully compatible with the ideal MHD system. Consequently, the solution of Eq. (12) is prone to include spurious $\nabla \cdot \vec{B}$ errors that are physically inadmissible.

To illustrate this further, we consider the update to the initial magnetic field data associated with a particular realization of the control variable, given by

$$\boldsymbol{B}^0 = \check{\boldsymbol{B}}^0 + \boldsymbol{L}_{\boldsymbol{B}}\boldsymbol{V}_{\boldsymbol{B}}\,, \tag{15}$$

where the $\boldsymbol{B}$ subscripts indicate a restriction to the magnetic field components, and $\check{\boldsymbol{B}}^0$ denotes the *a priori* discrete initial magnetic field data. The product, $\boldsymbol{L}_{\boldsymbol{B}}\boldsymbol{V}_{\boldsymbol{B}}$, is not divergence-free in general and therefore pollutes the magnetic field estimate with unphysical modes despite a divergence-free *a priori* estimate. Moreover, the resulting magnetic field may exhibit long-wavelength $\nabla \cdot \vec{B}$ errors, as opposed to the short-wavelength deviations typically induced by discretization. If not accounted for, these $\nabla \cdot \vec{B}$ errors can lead to numerically unstable solutions that can prevent practical simulations in many cases.

To ensure the variational DA analysis does not introduce non-zero $\nabla \cdot \vec{B}$ errors, the magnetic field is parametrized by the vector potential, $\vec{A}$, according to

$$\vec{B} = \nabla \times \vec{A}\,, \tag{16}$$

which follows from the Helmholtz decomposition theorem. In this form, any choice of $\vec{A}$ ensures a physically consistent, divergence-free, magnetic field. Accordingly, magnetic field increments, $\delta\vec{B}$, induced by DA are also associated with a corresponding increment in the magnetic vector potential via

$$\delta\vec{B} = \nabla \times \delta\vec{A}\,. \tag{17}$$

Discretization of Eq. (17) yields an analogous statement, given by

$$\boldsymbol{\delta B} = \boldsymbol{C}\boldsymbol{\delta A}\,, \tag{18}$$

under a suitable discrete curl operator, $\boldsymbol{C}$, to be defined later in this section.

The increment, $\boldsymbol{\delta B}$, is further decomposed via CVT to yield

$$\boldsymbol{\delta B} = \boldsymbol{C L_A V_A}\,, \tag{19}$$

where $\boldsymbol{L_A}$ is the Cholesky factor of the background-error covariance matrix for the magnetic vector potential, and $\boldsymbol{V_A}$ is the corresponding control variable associated with the initial magnetic vector potential data. Replacing the second term in the right hand side of Eq. (15) with Eq. (19) yields an update rule for the magnetic field, given by

$$\boldsymbol{B}^0 = \check{\boldsymbol{B}}^0 + \boldsymbol{C L_A V_A}\,. \tag{20}$$

In this form, any choice of $\boldsymbol{V_A}$ results in discretely divergence-free initial magnetic field data, provided the *a priori* estimate, $\check{\boldsymbol{B}}^0$, is divergence-free, and the curl operator is mimetic with respect to the discrete divergence operator of interest. In addition, this parametrization implicitly replaces the magnetic field entries of the error covariance matrix with

$$\Sigma_{\boldsymbol{B}} = \tilde{\boldsymbol{L}}_{\boldsymbol{B}} \tilde{\boldsymbol{L}}_{\boldsymbol{B}}^{\intercal}\,, \quad \text{with,} \quad \tilde{\boldsymbol{L}}_{\boldsymbol{B}} \equiv \boldsymbol{C L_A}\,. \tag{21}$$

By consequence of the magnetic vector potential formulation, the increment, $\boldsymbol{\delta A}$, generally contains redundant degrees of freedom unless further restricted. Redundancies arise because vector potential fields are invariant under gauge transformations,

$$\vec{A} \mapsto \vec{A} + \nabla\varphi\,, \tag{22}$$

for any scalar field φ. Different choices of $\vec{A}$ related by such transformations correspond to the same physical magnetic field, as implied by the identity $\nabla \times \nabla(\cdot) = \vec{0}$. To constrain the redundant degrees of freedom, while simultaneously improving computational efficiency, we employ the so-called axial gauge, while also enforcing suitable boundary conditions. This gauge choice forces the x component of the vector potential to vanish. Indeed, for an arbitrary vector potential field, a corresponding scalar potential can always be found so that $A_x = 0$. Although this gauge-choice is rather uncommon in the literature, it enables the representation of the vector potential with two components, as opposed to three, thereby reducing both memory requirements and the number of variables involved in the DA optimization problem. Moreover, under this gauge choice, the curl operator, $\boldsymbol{C}$, can be represented as an injective linear operator (i.e., of full column rank, with distinct 2D fields mapping to distinct 3D fields), a property which will be leveraged in the subsequent analysis.

Restricting the magnetic field increments to a divergence-free subspace necessarily limits the rank of the magnetic field background covariance matrix. This rank deficiency may be illustrated by examining Eq. (21), where the curl operator, if represented as a matrix, is of size $N_{\boldsymbol{B}} \times N_{\boldsymbol{A}}$, and the Cholesky factor, $L_{\boldsymbol{A}}$, is of size $N_{\boldsymbol{A}} \times N_{\boldsymbol{A}}$. Here, $N_{\boldsymbol{B}}$ and $N_{\boldsymbol{A}}$ denote the number of cell centered solution quantities associated with the magnetic field and the magnetic vector potential, respectively. Under the axial gauge, the number of vector potential quantities is reduced, so that $N_{\boldsymbol{A}} < N_{\boldsymbol{B}}$. It follows that the rank of the background covariance matrix, $\boldsymbol{\Sigma_B}$, is bounded by

$$\operatorname{rank}(\boldsymbol{\Sigma_B}) \leq N_{\boldsymbol{A}}\,, \tag{23}$$

and is therefore rank deficient. As a result, the variational DA cost function, J, given in Eq. (12), which involves the inverse of $\boldsymbol{\Sigma_B}$, becomes undefined. Nevertheless, the CVT form of the cost function, given by

$$J(\boldsymbol{V}) = \frac{1}{2}\sum_{n=1}^{N} ||\boldsymbol{\mathcal{H}}(\boldsymbol{\mathcal{U}}^n) - \boldsymbol{d}^n||^2_{\Sigma_d^{-1}} + \frac{1}{2}||\boldsymbol{V}||_2^2\,, \tag{24}$$

remains computable in practice. This naturally raises the question of what, exactly, is physically minimized under this new framing. As is proved below, Eq. (24) is physically meaningful if one replaces the inverse, $(\boldsymbol{\Sigma_B})^{-1}$, with the Moore-Penrose inverse[1], $(\boldsymbol{\Sigma_B})^{+}$, thereby providing a natural generalization of the DA problem. This equivalence is only guaranteed if the injectivity of the curl operator can be ascertained; otherwise, additional terms must be included in the background component of Eq. (24). As a result of this generalization, the implementation of the proposed divergence-free CVT requires only minimal modifications to standard variational DA formulation.

Proposition 1. *The control variable form of the DA cost function, when used with the magnetic vector potential CVT, is equivalent to the physical form, provided that the background covariance inverse is replaced by the corresponding Moore–Penrose inverse; that is,*

$$\frac{1}{2}\sum_{n=1}^{N}||\boldsymbol{\mathcal{H}}(\boldsymbol{\mathcal{U}}^n) - \boldsymbol{d}^n||^2_{\Sigma_d^{-1}} + \frac{1}{2}||\boldsymbol{\mathcal{U}}^0 - \check{\boldsymbol{\mathcal{U}}}^0||^2_{\Sigma^+_{\mathcal{U}}} = \frac{1}{2}\sum_{n=1}^{N}||\boldsymbol{\mathcal{H}}(\boldsymbol{\mathcal{U}}^n) - \boldsymbol{d}^n||^2_{\Sigma_d^{-1}} + \frac{1}{2}||\boldsymbol{V}||^2_2 \tag{25}$$

Proof. The observation terms are identical, so we focus on the background terms. Since the background covariance is block diagonal and the CVT only modifies magnetic components, it suffices to consider these, reducing the proposition to

$$\frac{1}{2}||\boldsymbol{B}^0 - \check{\boldsymbol{B}}^0||^2_{\Sigma^+_{\boldsymbol{B}}} = \frac{1}{2}||\boldsymbol{V}_{\boldsymbol{A}}||^2_2 \,. \tag{26}$$

Expanding the left-hand side using the background covariance definition yields

$$\text{LHS} = \boldsymbol{V}_{\boldsymbol{A}}^\intercal \tilde{\boldsymbol{L}}_{\boldsymbol{B}}^\intercal (\tilde{\boldsymbol{L}}_{\boldsymbol{B}} \tilde{\boldsymbol{L}}_{\boldsymbol{B}}^\intercal)^+ \tilde{\boldsymbol{L}}_{\boldsymbol{B}} \boldsymbol{V}_{\boldsymbol{A}} \,. \tag{27}$$

Applying the identity $\boldsymbol{M}^+ = \boldsymbol{M}^\intercal(\boldsymbol{M}\boldsymbol{M}^\intercal)^+$, results in

$$\text{LHS} = \boldsymbol{V}_{\boldsymbol{A}}^\intercal \tilde{\boldsymbol{L}}_{\boldsymbol{B}}^+ \tilde{\boldsymbol{L}}_{\boldsymbol{B}} \boldsymbol{V}_{\boldsymbol{A}} \,. \tag{28}$$

Finally, the right-hand side follows from the injectivity of $\boldsymbol{C}$, the invertibility of $\boldsymbol{L}_{\boldsymbol{A}}$, and the well known identity, $\boldsymbol{M}^+\boldsymbol{M} = \boldsymbol{I}$ for injective matrices. □

The specification of the curl operator is considered next. This operator should ideally be defined so that a suitable discretization of the solenoidal condition, $\nabla \cdot \vec{B} = 0$, is maintained to machine accuracy; we call this the mimetic property. In general, it is not possible to construct a curl operator that is mimetic under arbitrary divergence operators, so the choice of discretization requires care. Nevertheless, if the mimetic property cannot be enforced, many curl operators will yield magnetic fields whose divergence vanishes in the limit of grid refinement.

Here, the curl operator is defined using a standard finite-volume discretization, where integration is carried out over each computational cell, and Gauss' theorem is applied to interchange volume integrals with surface integrals. Specifically, the cell-averaged magnetic field associated with a vector potential is given by

$$\vec{B}_K = \frac{1}{\mathcal{V}_K}\int_K \nabla \times \vec{A}\mathrm{d}V \tag{29}$$

$$= \frac{1}{\mathcal{V}_K}\oint_{\partial K} \hat{n} \times \vec{A}\mathrm{d}S \tag{30}$$

$$\approx \frac{1}{\mathcal{V}_K}\sum_{q=1}^{N_q}(\hat{n}_q \mathcal{A}_q \times \vec{A}_q)_K \,, \tag{31}$$

where the vector potential at cell interfaces is taken as the arithmetic mean of the neighbouring cell-centred values. With this definition, and when applied on a Cartesian grid, the discrete curl operator is

[1]The Moore-Penrose inverse generalizes the notion of matrix inversion and applies to matrices that are not necessarily square or invertible.

mimetic with respect to the divergence operator defined by

$$\frac{1}{\mathcal{V}_K}\int_K \nabla\cdot\vec{B}\mathrm{d}V \approx \frac{1}{\mathcal{V}_K}\sum_{q=1}^{N_q}(\hat{n}_q\mathcal{A}_q\cdot\vec{B}_q)_K\,, \tag{32}$$

where the cell-interface magnetic field is again specified by arithmetic averaging. Similar mimetic operators could also be constructed on general curvilinear grids, but this lies beyond the scope of this paper. See, for example, Appendix B of [23].

3.3 Adjoint-Based Numerical Solution Method

The CVT form of the variational DA optimization problem reads

$$\begin{aligned} \underset{\boldsymbol{V}}{\arg\min} \quad & J = \frac{1}{2}\sum_{n=1}^{N}||\boldsymbol{\mathcal{H}}(\boldsymbol{\mathcal{U}}^n) - \boldsymbol{d}^n||^2_{\boldsymbol{\Sigma}_d^{-1}} + \frac{1}{2}||\boldsymbol{V}||^2_2\,, \\ \text{s.t.} \quad & \tilde{\boldsymbol{R}} = \boldsymbol{0}\,. \end{aligned} \tag{33}$$

This constrained optimization problem can more readily be solved by casting it as an unconstrained problem through straightforward application of the method of Lagrange multipliers. Adopting this approach, a Lagrangian, $\mathcal{L}$, is constructed according to

$$\mathcal{L}(\boldsymbol{V},\boldsymbol{U}^1,\cdots,\boldsymbol{U}^N,\boldsymbol{\lambda}^1,\cdots,\boldsymbol{\lambda}^N) = J + \sum_{n=0}^{N-1}(\boldsymbol{\lambda}^{n+1})^\intercal\tilde{\boldsymbol{R}}(\boldsymbol{U}^{n+1},\boldsymbol{U}^n)\,, \tag{34}$$

where the Lagrange multiplier, $\boldsymbol{\lambda}$, is introduced. It is noted that Lagrange multipliers are often also referred to as adjoint variables in the PDE constrained optimization literature—this convention is also adopted herein. The Lagrangian is minimized by finding the stationary points where the derivatives of $\mathcal{L}$ with respect to each of the solution variables, control variable, and adjoint variables vanish, which results in the so-called Euler-Lagrange equations, given by

$$\widetilde{\boldsymbol{R}}(\boldsymbol{U}^n,\boldsymbol{U}^{n-1}) = \boldsymbol{0}, \qquad \forall n\in\{1,\cdots,N\}\,, \tag{35}$$

$$\boldsymbol{\lambda}^N = -\left[\frac{\partial\boldsymbol{\mathcal{H}}^N}{\partial\boldsymbol{U}^N}\right]^\intercal\boldsymbol{\Sigma}_d^{-1}(\boldsymbol{\mathcal{H}}(\boldsymbol{U}^N)-\boldsymbol{d}^N)\,, \tag{36}$$

$$\boldsymbol{\lambda}^n = \left[\frac{\partial\boldsymbol{\mathcal{M}}^n}{\partial\boldsymbol{U}^n}\right]^\intercal\boldsymbol{\lambda}^{n+1} - \left[\frac{\partial\boldsymbol{\mathcal{H}}^n}{\partial\boldsymbol{U}^n}\right]^\intercal\boldsymbol{\Sigma}_d^{-1}(\boldsymbol{\mathcal{H}}(\boldsymbol{U}^n)-\boldsymbol{d}^n) \qquad \forall n\in\{1,\cdots,N-1\}\,, \tag{37}$$

$$\boldsymbol{L}^\intercal\boldsymbol{\lambda}^0 = \boldsymbol{V}\,. \tag{38}$$

where

$$\boldsymbol{\lambda}^0 \equiv \left[\frac{\partial\boldsymbol{\mathcal{M}}^0}{\partial\boldsymbol{U}^0}\right]^\intercal\boldsymbol{\lambda}^1\,, \tag{39}$$

and $\boldsymbol{\mathcal{M}}$ is the time marching operator given by the right hand side of Eq. (7).

The Euler-Lagrange equations encompass the discrete ideal MHD equations given in Eq. (35), the so-called adjoint equations given in Eq. (37), with initial and final conditions given in Eqs. (38) and (36) respectively. The adjoint equations form a two-point boundary problem that is in practice quite challenging to solve directly. A reduced-space approach is taken here where the Lagrangian is iteratively minimized by employing the gradient-based Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [24]. At each iteration, the ideal MHD equations are evolved forward in time by the explicit time marching scheme described in Section 2, this is followed by the backwards-in-time evolution of the adjoint variables by

Eqs. (36) and (37). The cost function gradient is then finally computed according to

$$\nabla_{\boldsymbol{V}} J = -\boldsymbol{L}^{\intercal} \boldsymbol{\lambda}^{0} + \boldsymbol{V} , \tag{40}$$

which is then used to update the initial data via the BFGS scheme. The numerical minimization scheme is iterated until Eq. (38) is approximately satisfied. The numerical evaluation of the various vector-Jacobian products contained in Eqs. (35)–(38) is exhaustively detailed in the Ph.D. thesis of the first author [9].

4. Comparison of Treatment Strategies for $\nabla \cdot \vec{B}$ Errors

Numerical experiments are considered in this section to test the magnetic vector potential formulation for variational DA proposed herein. A series of observing system simulation experiments (OSSEs) are conducted, where synthetic observations are assimilated into a background simulation with prescribed errors in attempts to recover the ground-truth state. The standard and vector potential formulations of variational DA are considered. In the former, Powell's approach [4] is the sole mechanism for removing non-zero $\nabla \cdot \vec{B}$ errors produced by assimilation or discretization. These strategies are hereafter referred to as the Powell and vector potential approaches.

OSSEs are a standard validation study for new data-assimilation systems that allow for a controlled analysis of the information extracted from observations. Here, each OSSE consists of a "ground truth" state, $\boldsymbol{U}_{\mathrm{t}}$, and a background state whose initial data is drawn from $\mathcal{N}(\boldsymbol{U}_{\mathrm{t}}^{0}, \Sigma_{\boldsymbol{U}}^{0})$. We note that in the absence of observational data, the background evolution is representative of a forward ideal MHD prediction. To ensure that the background magnetic field data is divergence-free, the covariance model employed in the vector potential approach is used for initiation. To ensure consistency between the Powell and vector potential approaches, covariance parameters for the magnetic vector potential components are chosen such that the resulting magnetic field perturbations reproduce the magnitudes arising from direct specification of the magnetic field background covariance.

Noisy synthetic measurements are generated by sparsely sampling from the ground-truth solution, i.e. drawing samples from $\mathcal{N}(\boldsymbol{\mathcal{H}}(\boldsymbol{U}_{\mathrm{t}}^{n}), \Sigma_{\boldsymbol{d}}^{n})$. This synthetic data is then assimilated, with the background initial conditions serving as the *a priori* data. Errors are quantified by examining the point-wise root-mean-square (RMS) error, given by

$$\epsilon_K = \frac{||\boldsymbol{W}_K - \boldsymbol{W}_{\mathrm{t},K}||_2}{\sqrt{8}} , \tag{41}$$

the total RMS error, given by

$$\epsilon = \frac{||\boldsymbol{W} - \boldsymbol{W}_{\mathrm{t}}||_2}{\sqrt{8 N_{\mathrm{cells}}}} , \tag{42}$$

where N_{cells} is the total number of cells in the mesh, and similarly the total error associated with $\nabla \cdot \vec{B}$ violations, given by

$$\epsilon^{\nabla \cdot \vec{B}} = \sqrt{\frac{1}{N_{\mathrm{cells}}} \sum (\nabla \cdot \vec{B}_K)^2} . \tag{43}$$

In all numerical experiments, an upper limit of 150 iterations of the L-BFGS optimizer was imposed as the convergence criterion.

4.1 Rotated Brio-Wu Shock Tube Problem

The canonical, MHD, shock-tube, problem of Brio and Wu [25] is considered here. A multidimensional extension of the Brio-Wu problem is employed, whereby the problem geometry is rotated by 45° so that waves travel in a multidimensional fashion in the direction normal to the initial discontinuity plane. After

Table 1: Imposed variance of synthetic observation noise in non-dimensional units associated with each of the conserved solution components employed in the rotated Brio-Wu shock tube OSSEs.

	ρ	ρu	ρv	ρw	B_x	B_y	B_z	e
$\sigma^2_{\text{obs.}} \times 10^5$	3	0.5	3	0.5	3	3	3	10

rotating, the ground-truth initial data is given by

$$\boldsymbol{W} = \begin{cases} \left[1, 0, 0, 0, -\dfrac{1}{4\sqrt{2}}, \dfrac{7}{4\sqrt{2}}, 0, 1\right]^{\mathsf{T}} & x + y \leq 0 \\ \left[\dfrac{1}{8}, 0, 0, 0, \dfrac{7}{4\sqrt{2}}, -\dfrac{1}{4\sqrt{2}}, 0, \dfrac{1}{10}\right]^{\mathsf{T}} & x + y > 0 \end{cases}. \tag{44}$$

This rotated Brio-Wu problem was studied by Powell [4] as a test case for the 8-wave ideal MHD formulation. Other studies (e.g., [23]) also employ this rotated geometry with differing initial shock-tube data.

Simulations are performed on a $800 \times 4 \times 4$ uniform grid with the x axis spanning $[-1, 1]$, and employing 192 CPU processors. Explicit time marching is applied for 0.2 units of simulation time, requiring approximately 1000 time steps. So-called shifted-extrapolation boundary conditions are applied at the positive and negative y-axis boundaries to permit uninterrupted wave motion in the rotated frame. Shifted-extrapolation boundary conditions are implemented by extrapolating data along $y + x =$ constant lines (as opposed $x =$ constant lines for constant extrapolation). Constant extrapolation boundary conditions are imposed in the remaining four boundaries.

Two types of observation systems are employed: a *dense* system consists of observations located along 10 equidistant slices of the x-axis, with observations in all of the 16 intercepted cells, and a *sparse* system with observations present on two slices located at $x = \pm 0.2$. In both cases, each observation measures all eight conserved solution quantities at its respective location, with a period of 0.001 time units. The variance of the imposed synthetic measurement noise is given in Table 1.

Background error covariances are specified via separable correlation functions with the stepwise covariance matrix decomposition method of Li *et al.* [26]. One-dimensional second-order autoregressive correlation functions are used in each Cartesian direction, with length scales and diagonal variances specified in Table 2. The covariance parameters of the magnetic vector potential components were chosen so that the resulting magnetic field perturbations match the magnitude of those produced with the direct modelling of the magnetic field background covariance.

The convergence history of the DA cost function, J, is shown in Fig. 1 for the four OSSEs considered here. The vector potential and Powell approach show similar convergence rates, with the vector potential strategy achieving modestly lower final cost function values with both the dense and sparse observation systems. An approximately constant convergence rate for the first ~ 25 optimization iterations is observed before slowing down.

Profiles of the mass density and RMS error are shown in Fig. 2, with data extracted from the $y = z = 0$ line. Predictions from the background and assimilated solutions are compared against the ground-truth state at the final simulation time ($t = 0.2$). The background initial data perturbation can be seen to induce significant errors in the solution wave speeds, resulting in large discrepancies in the location of the fast rarefaction wave head, the compound wave, the contact discontinuity, and the slow magnetosonic shock. Additionally, a local density amplification is observed at $x \approx 0.4$. The data assimilated solutions, in contrast, demonstrate accurate wave speeds and dampening of local perturbations. Quantitatively, all OSSEs exhibit significant reductions in the total RMS error compared to the background simulation, as summarized in Table 3. For the dense observation system, the Powell and vector potential approach yield comparable RMS error reduction factors, with the Powell approach showing marginally superior

Table 2: Length scales, $D_{(\cdot)}$, and diagonal variances, σ^2, employed in the background error covariance model based on second-order autoregressive correlation functions for the rotated Brio-Wu shock tube OSSEs.

Parameter	$\nabla \cdot \vec{B}$ Strategy	ρ	u_x	u_y	u_z	B_x	B_y	B_z	p	A_y	A_z
$\sigma^2 \times 10^3$	Powell	2	2	2	2	2	2	2	2	—	—
	Vector Potential	2	2	2	2	—	—	—	2	2	2
$D_x \times 10$	Powell	2	2	2	2	2	2	2	2	—	—
	Vector Potential	2	2	2	2	—	—	—	2	8	8
$D_y \times 10$	Powell	2	2	2	2	2	2	2	2	—	—
	Vector Potential	2	2	2	2	—	—	—	2	10	10
$D_z \times 10$	Powell	2	2	2	2	2	2	2	2	—	—
	Vector Potential	2	2	2	2	—	—	—	2	2	2

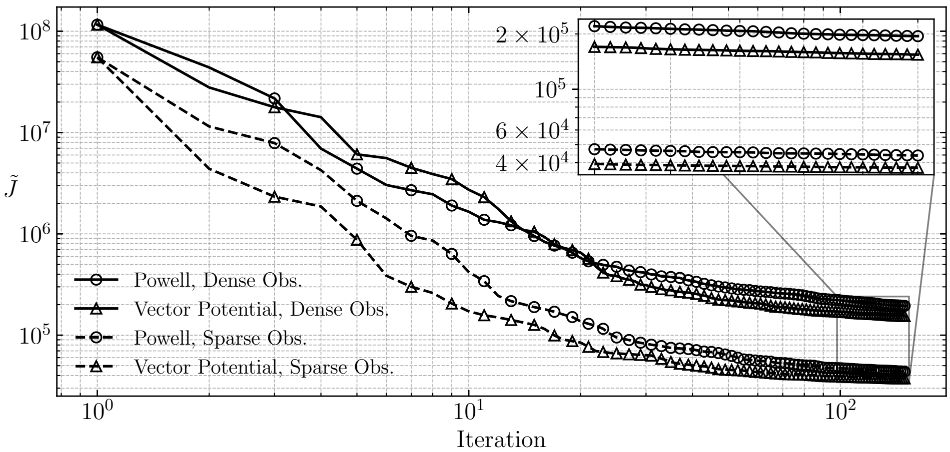

Figure 1: Convergence history of the DA cost function for the rotated Brio-Wu shock tube OSSEs. A close-up of the last 50 iterations is shown in the top right corner.

performance. Experiments involving the sparse observation system, in contrast, show markedly differing solutions. The vector potential approach achieves lower RMS errors as compared to the assimilation employing the Powell approach. Additionally, the vector potential approach faithfully reconstructs the density jump conditions across the slow magnetosonic shock front ($x \approx 0.4$), whereas the Powell approach does not. In both cases, the RMS error is lowest in the proximity of observations and gradually relaxes to background levels near the boundaries of the domain where observations have the least impact.

To examine the effects of DA on the magnetic field divergence-free property, the component of the magnetic field that lies parallel to wave motion, $B_{\|}$, is shown in Fig. 3 along with $\nabla \cdot \vec{B}$. These profiles are taken from the $y = z = 0$ line at the final time step. We note that in the absence of perturbations and discretization errors, $B_{\|}$ takes a constant value of 0.75 across the shock tube. Additionally, the total RMS error of the $\nabla \cdot \vec{B}$ condition is reported in Table 3. These results suggest that given sufficiently dense observations, the Powell approach can suppress the $\nabla \cdot \vec{B}$ errors induced by DA. However, when sparse observations are assimilated, substantial $\nabla \cdot \vec{B}$ errors can be introduced, as evident in the top right panel of Fig. 3 at $x \approx 0.4$, which give rise to the incorrect jump conditions observed in Fig. 2. The vector potential approach, on the other hand, does not produce any $\nabla \cdot \vec{B}$ errors associated with DA; all of the $\nabla \cdot \vec{B}$ fluctuations are due to discretization error. As a result, $\nabla \cdot \vec{B}$ errors remain small (Table 3), even in the case with sparse observational data, and $B_{\|}$ is well assimilated (Fig. 3, bottom row).

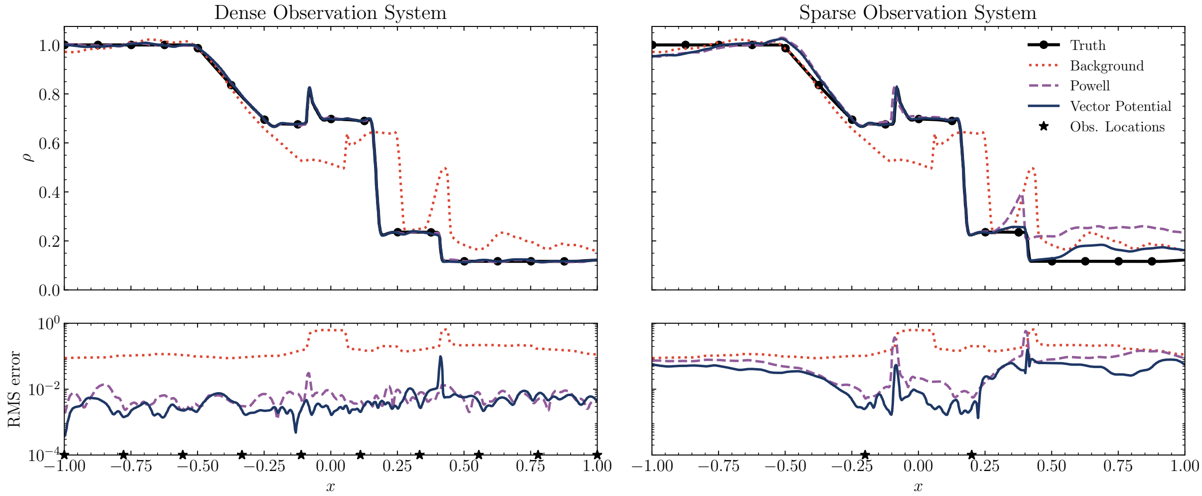

Figure 2: Profiles of the rotated Brio-Wu shock tube mass density (top row) predicted by the ground-truth, background, and data assimilated solutions with synthetic data arising from the dense (left) and sparse (right) observation systems. The RMS error of all primitive quantities is shown in each case on the bottom row. In all panels, the data is extracted from the three-dimensional simulation domain along the $y = z = 0$ line at the final time step of $t = 0.2$. Additionally, star symbols are used to denote the location of observations along the x-axis.

4.2 Magnetically Confined Thermal Expansion Initial-Value Problem

The second unsteady initial-value problem (IVP) considered here for DA involves a thermal expansion process, confined by magnetic field lines in a Z-pinch configuration. This IVP leads to hyperbolic non-linear wave transport with different components of the solution propagating at rather widely different velocities. The thermodynamic initial data was first formulated by Prof. Bram van Leer at the University of Michigan as a simple test case for gas dynamics simulations. The initial velocity field is purely radial, and is given by

$$u_r = \begin{cases} 0 & 0 \le r < \frac{1}{2}\,, \\ \frac{1}{\gamma}\left(1 + \tanh\left(\frac{r-1}{\frac{1}{4} - (r-1)^2}\right)\right) & \frac{1}{2} < r \le \frac{3}{2}\,, \\ \frac{2}{\gamma} & r \ge \frac{3}{2} \end{cases}, \tag{45}$$

Table 3: Root-mean-square errors associated with the primitive solution vector and the magnetic field divergence at the final simulation time step ($t = 0.2$) for the rotated Brio-Wu shock tube problem. Both absolute and relative percentage errors are reported.

Observation System	Case	ϵ	$\frac{\epsilon}{\epsilon_{\text{backg.}}}$	$\epsilon^{\nabla\cdot\vec{B}}$	$\frac{\epsilon^{\nabla\cdot\vec{B}}}{\epsilon^{\nabla\cdot\vec{B}}_{\text{backg.}}}$
—	Background	0.229	100%	0.703	100%
Dense	Powell	0.007	2.91%	0.441	62.7%
	Vector potential	0.008	3.31%	0.235	33.5%
Sparse	Powell	0.090	39.4%	0.701	99.7%
	Vector potential	0.042	18.3%	0.406	57.8%

[i] The “backg.” subscript indicates the error of the background simulation.

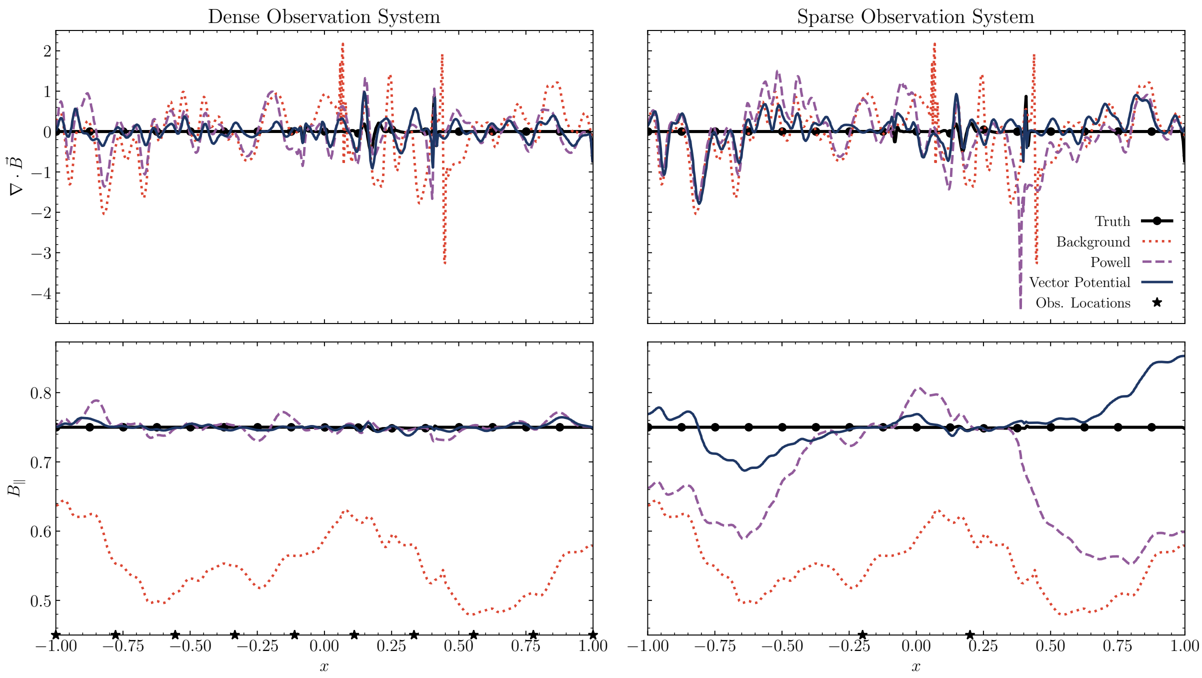

Figure 3: Profiles of the rotated Brio-Wu shock tube $\nabla \cdot \vec{B}$ errors (top row) predicted by the ground-truth, background, and data assimilated solutions with synthetic data arising from the dense (left) and sparse (right) observation systems. The parallel component of the magnetic field is shown in each case on the bottom row. In all panels, the data is extracted from the three-dimensional simulation domain along the $y = z = 0$ line at the final time step of $t = 0.2$. Additionally, star symbols are used to denote the location of observations along the x-axis.

with $\gamma = 2$. The remaining thermodynamic quantities are specified by letting both the entropy and the radial Riemann invariant be constant, and by prescribing $\rho = \gamma$, $a = p = 1$, at the origin, resulting in

$$u_r + \frac{2}{\gamma - 1} a = \frac{2}{\gamma - 1} , \tag{46}$$

$$\rho = \gamma a^{\frac{2}{\gamma-1}} , \tag{47}$$

$$p = \frac{\rho a^2}{\gamma} , \tag{48}$$

where a is the sound speed.

This test case is extended here to conducting plasmas by requiring that the Alfvén speed, c_a, be initially identical to the sound speed,

$$c_a \equiv \sqrt{\frac{|\vec{B}|^2}{\rho}} = a , \tag{49}$$

and by considering a purely azimuthal magnetic field so as to maintain solenoidality. This magnetic field configuration yields magnetic forces that initially confine the plasma inwards due to the predominance of magnetic tension over plasma pressure-gradient forces. As the flow evolves, the dynamic balance of outward and inward forces leads to both compressive and expansive motion throughout different regions of the spatial-temporal domain.

The IVP is simulated in a nearly two-dimensional $[-6, 6] \times [-6, 6]$ square domain with 400 computational cells in both the x and y directions. The third spatial dimension is included for conformity with the three-dimensional computer code implementation. The z direction is thus discretized with four cells over a

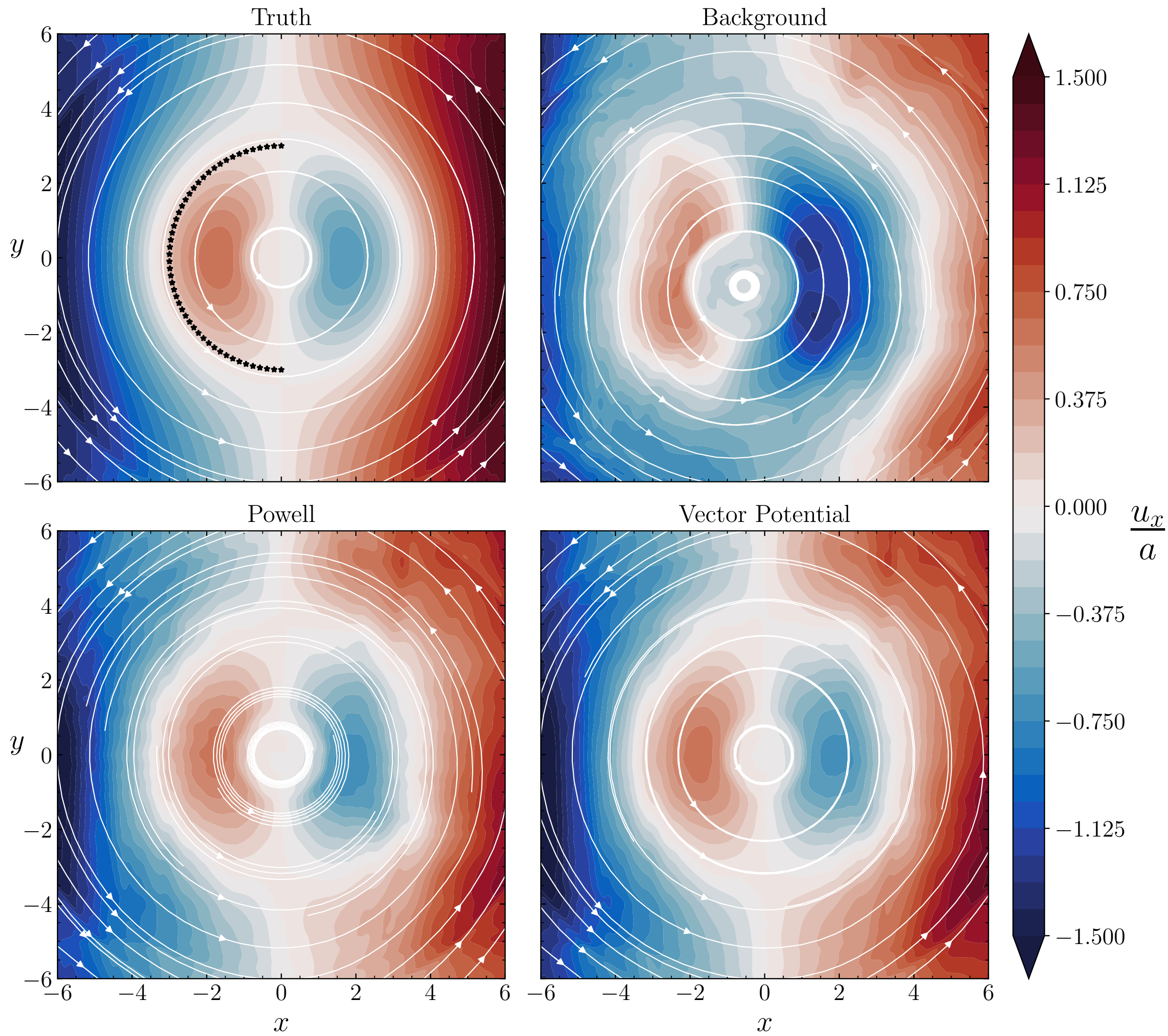

Figure 4: Snapshots of the flow-fields associated with the magnetically confined thermal expansion problem predicted by the ground-truth, background, and data assimilated solutions at the final simulation time of $t = 5$. Contours of the x-directed Mach number are shown along with magnetic field lines. The top-left panel also shows the locations of the synthetic observers used in the OSSEs, indicated by black star symbols.

distance of 0.12 units to maintain a homogeneous grid structure. The solution is evolved from 0 to 5 units in time, allowing for waves to exit the domain, and for the compression of the central core. Supersonic outflow boundary conditions are applied at all physical boundaries.

An observation system involving 50 synthetic sensors along the path of a semicircular arch of radius 3 is considered here. The observation locations are shown in the top left panel of Fig. 4. Observations of all eight conserved solution quantities are taken across 150 temporal snapshots. The imposed synthetic observation noise variances are listed in Table 4. Background error covariances are specified using separable correlation functions, with parameter values summarized in Table 5.

Figure 5 illustrates the convergence of the DA cost function, J, associated with each of the two OSSEs examined here. The Powell and vector potential formulations display broadly similar convergence behavior, although the vector potential approach yields marginally lower final cost function values. Convergence proceeds at an approximately uniform rate over the first ~ 40 iterations, followed by slower but also constant convergence rates.

Table 4: Imposed variance of synthetic observation noise associated with each of the conserved solution components employed in the magnetically confined thermal expansion OSSEs.

	ρ	ρu	ρv	ρw	B_x	B_y	B_z	e
$\sigma^2_{\text{obs.}} \times 10^6$	9	9	9	9	4	4	4	4

Table 5: Length scales, $D_{(\cdot)}$, and diagonal variances, σ^2, employed in the background error covariance model based on second-order autoregressive correlation functions for the magnetically confined thermal expansion OSSEs.

Parameter	$\nabla \cdot \vec{B}$ Strategy	ρ	u_x	u_y	u_z	B_x	B_y	B_z	p	A_y	A_z
$\sigma^2 \times 10^3$	Powell	8	8	8	8	4	4	1	0.125	—	—
	Vector Potential	8	8	8	8	—	—	—	0.125	2	2
D_x	Powell	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	—	—
	Vector Potential	1.2	1.2	1.2	1.2	—	—	—	1.2	6	6
D_y	Powell	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	—	—
	Vector Potential	1.2	1.2	1.2	1.2	—	—	—	1.2	6	6
D_z	Powell	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	—	—
	Vector Potential	1.2	1.2	1.2	1.2	—	—	—	1.2	10	10

Two-dimensional slices of the flow field are illustrated in Fig. 4, showing the x-directed Mach number and magnetic field lines predicted by the ground-truth, background, and assimilated solutions at the final time step ($t = 5$). Both the Powell and vector potential formulations yield close qualitative agreement with the ground truth data, correcting the non-circular field lines and the drift of the central core observed in the background simulation. Surprisingly, the right side of the solution domain is substantially corrected by DA, despite all observations being confined to the left half of the domain. This effect may be attributed to the adjoint-based approach employed here, which allows sensitivity information to be transported along flow characteristics. Nevertheless, high frequency perturbations remain evident in the assimilated solutions—most notably in the Powell formulation, which additionally exhibits spurious inward spiralling of field lines due to $\nabla \cdot \vec{B} = 0$ violations.

The combined point-wise RMS error of all primitive quantities is shown in Fig. 6, and the total RMS error over the solution domain is tabulated in Table 6. These results provide quantitative measures of the substantial amount of information extracted from the noisy, sparse, synthetic data. Additionally, the vector potential approach is once again shown to yield moderately improved performance over the Powell formulation.

Table 6: Total RMS errors associated with the primitive solution vector and the magnetic field divergence at the final simulation time step ($t = 5$) for the magnetically confined thermal expansion OSSEs. Both absolute and relative percentage errors are reported.

Case	ϵ	$\dfrac{\epsilon}{\epsilon_{\text{backg.}}}$	$\epsilon^{\nabla\cdot\vec{B}}$	$\dfrac{\epsilon^{\nabla\cdot\vec{B}}}{\epsilon^{\nabla\cdot\vec{B}}_{\text{backg.}}}$
Background	0.0896	100%	0.0174	100%
Powell	0.0247	27.5%	0.0144	82.9%
Vector potential	0.0193	21.6%	0.0140	80.3%

[i] The "backg." subscript indicates the error of the background simulation.

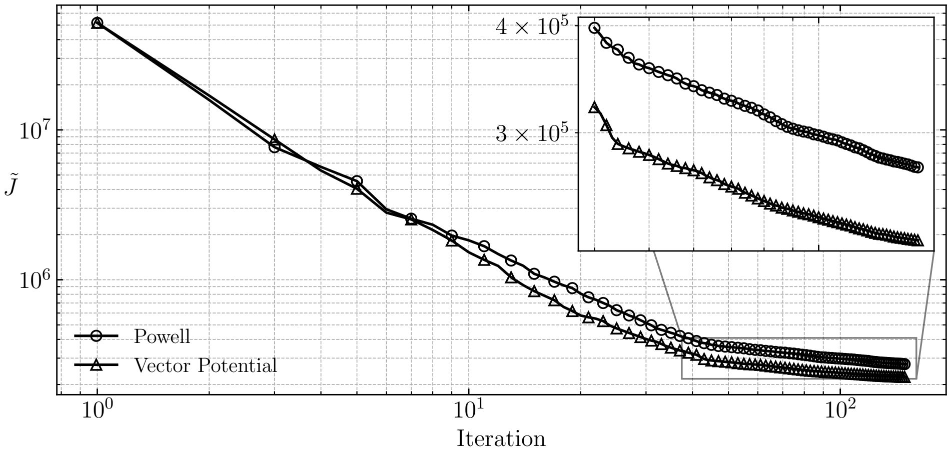

Figure 5: Convergence history of the DA cost function for the magnetically confined thermal expansion OSSEs. A close-up of the last 120 iterations is shown in the top right corner.

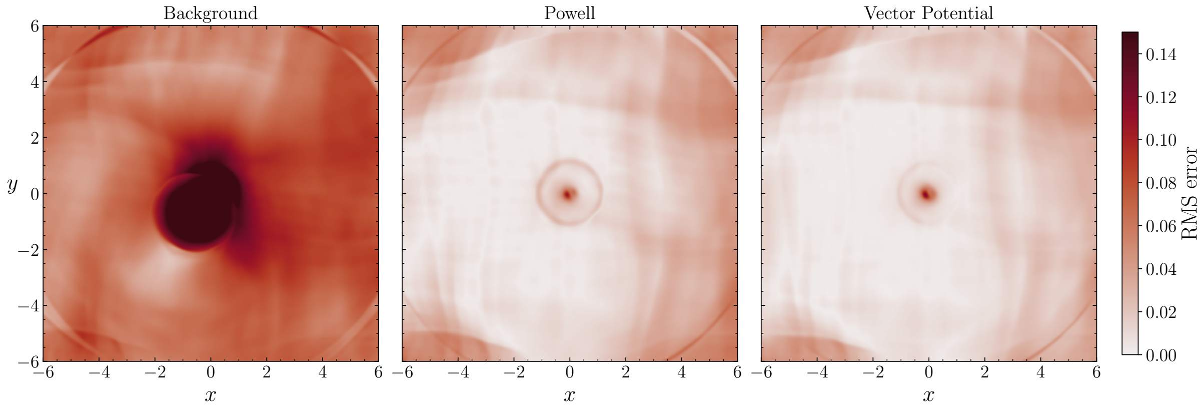

Figure 6: Snapshots of the point-wise RMS error associated with the magnetically confined thermal expansion problem pertaining to the background and data assimilated solutions at the final simulation time of $t = 5$.

The presence of non-zero $\nabla \cdot \vec{B}$ errors in the background and assimilated solutions is shown in Fig. 7, both for the initial and final solution time steps. All configurations exhibit violations to the divergence-free property. For the background and vector potential formulation, these errors are due to discretization, whereas the Powell approach incurs additional non-physical magnetic field modes induced by non divergence-preserving DA increments. These additional errors are introduced through the initial data and then propagate throughout the solution domain. As can be seen in the lower middle panel of Fig. 7, these errors persist to the end of the simulation window, meaning that the Powell formulation does not remove $\nabla \cdot \vec{B}$ errors sufficiently fast in this case. Despite these differences, the total $\nabla \cdot \vec{B}$ error present in the Powell approach is comparable to that of the vector potential approach, as is shown in Table 6. This occurs because the divergence-free violations associated with discretization are significantly larger, and therefore dominate, in comparison to those instilled by DA. Therefore, regions polluted by large $\nabla \cdot \vec{B}$ discretization errors do not benefit substantially more from the vector potential approach over the baseline Powell formulation. Conversely, in regions with smaller $\nabla \cdot \vec{B}$ discretization errors, the vector potential approach yields noticeable improvements over the Powell approach.

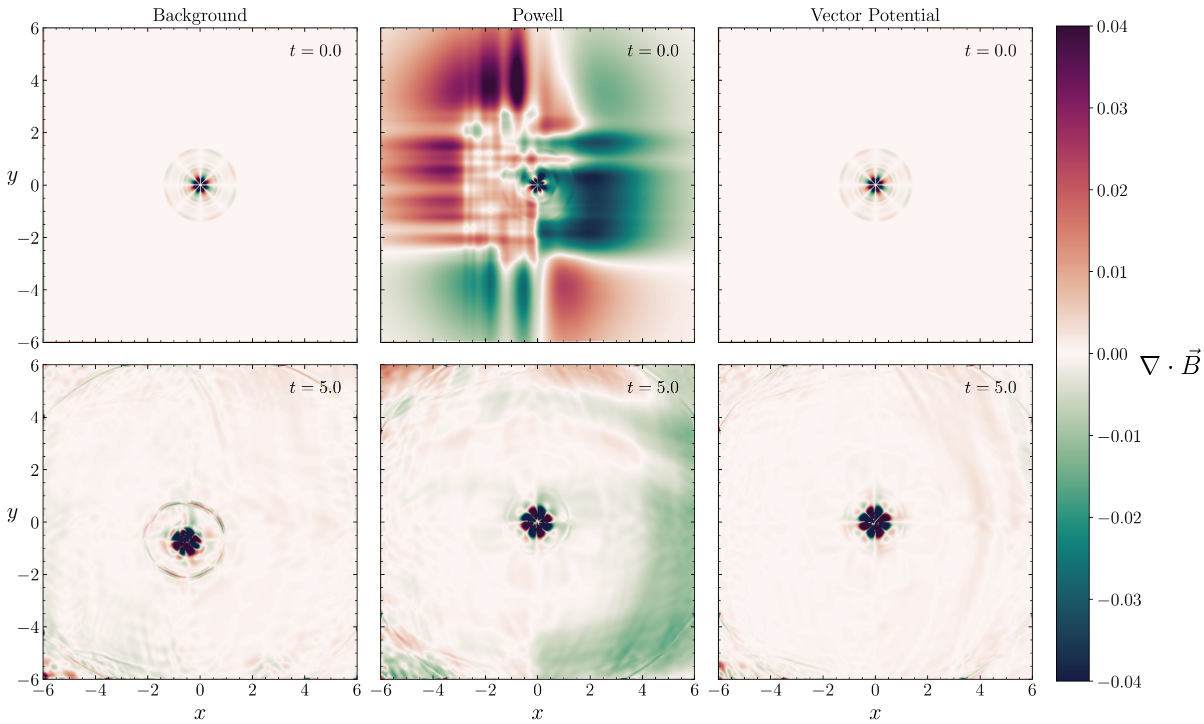

Figure 7: Snapshots of the $\nabla \cdot \vec{B}$ error associated with the magnetically confined thermal expansion problem pertaining to the background and data assimilated solutions at the initial (top row) and final (bottom row) simulation times.

5. Summary and Conclusions

This paper proposes and numerically evaluates a new method for enforcing the divergence-free condition of the magnetic field in the context of MHD-based variational DA. While the new method does not aim to modify the divergence-preserving properties of the underlying MHD scheme, it ensures that DA-based increments of magnetic field data are divergence-free. This new method is based on an implicit magnetic vector potential representation of solution errors, yet it does not require *a priori* data to be prescribed via a vector potential. This avoids the inversion (or "uncurling") of $\vec{B} = \nabla \times \vec{A}$ which would be necessary otherwise. However, magnetic covariances are expressed in terms of correlations in the vector potential, which may prove more challenging in practical applications.

Observing system simulation experiments were conducted based on two highly non-linear initial-value problems. The first involves a rotated version of Brio and Wu's shock tube problem [25, 4], while the second considers a magnetically confined thermal expansion process. A total of six configurations were tested, with all demonstrating substantial improvements of the solution accuracy by the assimilation of synthetic noisy data. The most significant error reduction occurred during the early stages of the numerical minimization, where convergence proceeded at an approximately constant rate.

The main benefit of the vector potential approach is that it does not produce any $\nabla \cdot \vec{B}$ errors through assimilation, whereas the baseline Powell approach [4] can only remove such errors after they are introduced. As a result, the errors associated with $\nabla \cdot \vec{B}$ violations were considerably smaller with the vector potential approach in all cases. The vector potential approach also yielded comparable RMS error reduction to the Powell approach for dense observation systems, and markedly lower errors for sparse observation systems. The comparable performance for dense observations is expected, since DA produces only small $\nabla \cdot \vec{B}$ errors when observations are abundant, so there is little for the vector potential approach to correct. The improvements are largest for sparse observations, where the $\nabla \cdot \vec{B}$ errors produced by DA are otherwise most significant. In turn, these improvements lead to better predictions of flow features such as the jump conditions across a slow magnetosonic shock front and the avoidance of spurious radial components in purely azimuthal magnetic field lines. However, we also observed flow regions where discretization-induced $\nabla \cdot \vec{B}$ errors dominated over those introduced by DA with the Powell approach. In these regions, the negative effects of $\nabla \cdot \vec{B}$ errors cannot be alleviated by the proposed method. Consequently, a combination of the proposed approach with a divergence-free preserving ideal MHD scheme, such as the constrained transport method of Evans and Hawley [7], represents a promising direction for future research.

References

[1] N. M. Narechania, L. Nikolić, L. Freret, H. De Sterck, and C. P. T. Groth. An integrated data-driven solar wind — CME numerical framework for space weather forecasting. *Journal of Space Weather and Space Climate*, 11(8), 2021.

[2] K. Li, A. Jackson, and P. W. Livermore. Variational data assimilation for a forced, inertia-free magnetohydrodynamic dynamo model. *Geophysical Journal International*, 199(3):1662–1676, 2014.

[3] J. P. Freidberg. Ideal magnetohydrodynamic theory of magnetic fusion systems. *Reviews of Modern Physics*, 54(3):801, 1982.

[4] K. G. Powell. An approximate Riemann solver for magnetohydrodynamics (that works in more than one dimension). Report 94-24, ICASE, July 1994.

[5] K. G. Powell, P. L. Roe, T. J. Linde, T. I. Gombosi, and D. L. De Zeeuw. A solution-adaptive upwind scheme for ideal magnetohydrodynamics. *Journal of Computational Physics*, 154:284–309, 1999.

[6] J. U. Brackbill and D. C. Barnes. The effect of nonzero $\nabla \cdot$ B on the numerical solution of the magnetohydrodynamic equations. *Journal of Computational Physics*, 35:426–430, 1980.

[7] C. R. Evans and J. F. Hawley. Simulation of magnetohydrodynamic flows: A constrained transport method. *Astrophysical Journal*, 332:659–677, 1988.

[8] J. H. Arnal and C. P. T. Groth. Quantitative comparison of variational and sequential data assimilation techniques for one-dimensional initial-value problems of ideal MHD. *Computer and Fluids*, 282:106373, 2024.

[9] J. H. Arnal. *Adjoint-Based Data Assimilation for Enhanced Magnetohydrodynamic Modelling of Space Plasmas*. PhD thesis, University of Toronto, June 2026.

[10] M. Lang and M. J. Owens. A variational approach to data assimilation in the solar wind. *Space Weather*, 17(1):59–83, 2019.

[11] L. Ivan, H. De Sterck, S. A. Northrup, and C. P. T. Groth. Three-dimensional MHD on cubed-sphere grids: Parallel solution-adaptive simulation framework. Paper 2011-3382, AIAA, June 2011.

[12] A. Harten, P. D. Lax, and B. van Leer. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM Review*, 25(1):35–61, 1983.

[13] T. J. Barth and D. C. Jespersen. The design and application of upwind schemes on unstructured meshes. Paper 89-0366, AIAA, January 1989.

[14] V. Venkatakrishnan. On the accuracy of limiters and convergence to steady state solutions. Paper 93-0880, AIAA, January 1993.

[15] C.-W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *Journal of Computational Physics*, 77:439–471, 1988.

[16] S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Review*, 43(1):89–112, 2001.

[17] R. Courant and K. O. Friedrichs. *Supersonic Flow and Shock Waves*. Springer, New York, 5th edition, 1999.

[18] S. K. Godunov. Symmetric form of the equations for magnetohydrodynamics. Report, Computer Center, Siberian Branch of USSR Academy of Sciences, 1972.

[19] S. K. Godunov. An interesting class of quasilinear systems. *Soviet Mathematics Doklady*, 2:947–949, 1961.

[20] A. Harten. High resolution schemes for hyperbolic conservation laws. *Journal of Computational Physics*, 49:357–393, 1983.

[21] A. C. Lorenc. Optimal nonlinear objective analysis. *Quarterly Journal of the Royal Meteorological Society*, 114(479):205–240, 1988.

[22] P. Courtier. Dual formulation of four-dimensional variational assimilation. *Quarterly Journal of the Royal Meteorological Society*, 123(544):2449–2461, 1997.

[23] G. Tóth. The $\nabla\cdot$ B=0 constraint in shock-capturing magnetohydrodynamics codes. *Journal of Computational Physics*, 161:605–652, 2000.

[24] D. C. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45(1-3):503–528, 1989.

[25] M. Brio and C. C. Wu. An upwind differencing scheme for the equations of ideal magnetohydrodynamics. *Journal of Computational Physics*, 75:400–422, 1988.

[26] D. Q. Li, T. Xiao, and Z. J. Cao L. M. Zhang. Stepwise covariance matrix decomposition for efficient simulation of multivariate large-scale three-dimensional random fields. *Applied Mathematical Modelling*, 68:169–181, 2019.