Quantitative comparison of variational and sequential data assimilation techniques for one-dimensional initial-value problems of ideal MHD

J.H. Arnal *, C.P.T. Groth

Institute for Aerospace Studies, University of Toronto, 4925 Dufferin Street, Toronto, ON M3H5T6, Canada

ARTICLE INFO

ABSTRACT

State-of-the-art predictions of the solar-wind and space weather phenomena are today largely based on the equations of magnetohydrodynamics (MHD). Despite their sophistication and success, the forecasting potential of global MHD models is often undermined by uncertainties in model inputs; the initial and boundary conditions are generally not known and must be estimated. This study therefore investigates the use of data assimilation strategies to minimize forecast errors in the context of initial-value problems of the one-dimensional ideal MHD equations. Several canonical MHD wave propagation problems involving both smooth and discontinuous solutions, including those having strongly non-linear behaviour with shocks, are considered in a set of twin experiments with varying synthetic observational data sparsity. Two data assimilation strategies are quantitatively compared, namely the Ensemble Kalman Filter (EnKF) and strong-constraint variational data assimilation. For the latter, the necessary adjoint model is derived, summarized, and validated. The study represents the first use of variational data assimilation applied to ideal magnetohydrodynamics and demonstrates its potential advantages over sequential approaches. In particular, for the numerical experiments considered herein, it is found that the variational approach consistently achieved superior performance and stability compared to the EnKF method. In addition, two different strategies for mitigating data assimilation induced errors associated with violation of the divergence-free property of the magnetic field are introduced and assessed. Finally, the present study provides the technical background and quantitative justification for future investigations of variational data assimilation aimed at enhancing three-dimensional simulations of the solar wind and space weather processes.

1. Introduction and background

The term "space weather" was first coined in 1959 by Thomas Gold of Harvard College Observatory [1]. It refers to conditions on the Sun and in the solar wind, and in the geospace environment of the magnetosphere, ionosphere, and thermosphere that can detrimentally influence the performance and reliability of space-borne and ground-based systems and adversely affect human life and health [2]. The harmful effects of space weather include disruptions to ground-based electric power transmission grids; increased corrosion of pipeline systems for oil, natural gas, and water; disruptions to satellite operations in near-earth orbit; disruptions to communication and navigation systems; and radiation hazards associated with high-energy particles for astronauts and aircraft crew at high altitudes. Advances in understanding of space weather are therefore a current high priority. Moreover, there is also high demand for accurate and reliable space weather forecasting capabilities so as to inform mitigation strategies in a manner analogous to that currently possible in the atmospheric weather and climate science communities [3].

Over the last several decades, a tremendous amount of effort has been dedicated to the computational modelling of space plasmas flows, with applications to heliospheric physics, the solar wind, and space weather science and forecasting. As with efforts in the atmospheric weather prediction and climate modelling communities, the development and use of improved and more accurate computational models and tools for space weather prediction has been the primary focus of solar wind and space weather related research effort. The space weather forecast models that have resulted from this effort are usually based on the equations of ideal magnetohydrodynamics (MHD) and make use of accurate and robust discretization methods for plasma flows with shocks, automatic mesh refinement, and impose boundary conditions derived from magnetic field measurements of the solar surface. The research related to space weather has focused on a range of issues including accurate discretization strategies, treatment of the divergence-free condition for the magnetic field, and gaining fundamental understanding of various physical phenomena associated with unsteady solar wind disturbances. While the research in this area is

---

far too extensive to review here, some examples of space weather forecasting tools based on global MHD modelling include the so-called Space Weather Modeling Framework (SWMF) developed in a series of studies by [4–12], ENLIL, a fully three-dimensional and time-dependent MHD model of the heliosphere developed by Odstrcil et al. [13–15], the EUropean Heliospheric FORecasting Information Asset (EUHFORIA) MHD-based simulation code [16,17], and the recent integrated data-driven solar wind, coronal mass ejection (CME), numerical framework for space weather forecasting of Narechania *at al.* [18]. Note that the SWMF model was used previously to provide one of the first global MHD simulations of a complete fully three-dimensional space weather event, spanning the initiation of a solar wind disturbance at the Sun's surface to its interaction with the Earth's magnetosphere [5].

Despite the research efforts briefly outlined above related to improved computational models and tools for space weather prediction, as well as the high level of sophistication and demonstrated efficacy of the resulting space weather forecasting tools, the predictive potential of global MHD-based models are currently not fully realized. In particular, space weather forecasts often fail to match accurately measurements and/or observations. This is largely due to poorly known values for a range of model input parameters. Suspected important sources of the forecast uncertainty include relatively large uncertainties in initial conditions, boundary data, and various sub-physics model parameters. It is noted that modern atmospheric weather and climate forecasts do not however rely solely on accurate numerical models and solution methods alone. They also make extensive use of data assimilation (DA) techniques, which combine real-time measurements and observational data from various sources with the predictions of the simulation models, effectively using the observational data to constrain the model predictions, so as to achieve more accurate and reliable forecasts [19–21]. DA is now a rather mature discipline in meteorological [19,21] and oceanographic [20,21] applications. Moreover, it is recognized as core to enabling today's atmospheric weather and climate forecasts. In contrast, the space weather forecasting community has only just begun to explore the full potential of such data-driven methods [22,23].

As introduced above, DA is the mathematical discipline of combining model outputs and measured data. DA strategies consider both the observations and model predictions as arising from underlying stochastic processes and crucial to these techniques is the appropriate weighting of both the observational and forecast model data to account for their uncertainty. One of two categories of approaches are usually adopted in assimilation methods: sequential or variational DA. In sequential DA, the model state is successively updated as observations become available. In the variational approach, measurements collected over a time window are used simultaneously to search for the corresponding optimal initial conditions. Examples of sequential data assimilation schemes include the Kalman filter (KF) [24], ensemble Kalman filter (EnKF) [25,26], and their variants [27,28]. An advantage of sequential methods is that they can often be implemented in a black-box fashion where the data assimilation algorithm is agnostic to the choice of model. Additionally, efficient parallel implementation of ensemble-based methods can be rather straight-forward. Variational methods [29–31], on the other hand, require the development of a model-specific adjoint. Moreover, efficient parallelizations are not readily achieved. The corresponding adjoint models of variational approaches are also often difficult to replicate and extend as they are often derived directly from the discrete representation of the forecast model, as opposed to a more general analytical formulation. Despite these challenges, better performance is often reported with variational DA techniques [32]. Additionally, variational methods allow the mapping of information gained from observations back in time when correcting forecasts [33]. A more complete review of the state-of-the-art of data assimilation is provided by Carrassi et al. [34] where the relative merits of various data assimilation strategies are discussed.

Several authors have previously conducted intercomparison studies between the EnKF and so-called Four-Dimensional Variational (4DVar)

methods. The EnKF approach was first examined as a possible substitute to 4DVar for numerical weather prediction (NWP) by Lorenc [35], who reported similar performance between both methods in medium-range NWP systems. Additionally, high-resolution satellite data was suspected to be more effectively assimilated by a 4DVar approach. Caya et al. [36] compared the EnKF and 4DVar method in the assimilation of radar data with a cloud-resolving model. Generally, better performance was observed with 4DVar for short time periods. In contrast, over several assimilation cycles, the EnKF yielded more accurate analyses. Kalnay et al. [37] also performed comparative studies of the EnKF and 4DVar for the Lorenz 1963 model and a quasi-geostrophic channel model. For both models, the EnKF method achieved either better or similar performance compared to 4DVar approach over short and long assimilation windows, respectively. However, when infrequent observations were assimilated, the 4DVar method obtained higher predictive accuracy. Several, more recent, studies comparing the EnKF and 4DVar approaches in over various disciplines have also been conducted. For the sake of brevity, however, they are not discussed in detail here. Nevertheless, from an examination of the literature pertaining to comparative studies of the EnKF and 4DVar methods it would appear that the preferred or optimal choice of the two strategies is not obvious and highly dependent on the numerical model employed, the type and availability of observations and measurement data, and the treatment of model uncertainties. For instance, some authors find relatively similar performance between both data assimilation methods [38–40], while others report significant differences and improvements depending on the method and application [41–46]. With this perspective in mind, it is evident that intercomparison studies of competing data assimilation methods can provide valuable insights when introducing data assimilation to novel areas such as the space-physics applications considered herein.

The application of DA strategies to forecast models based on MHD and related equations is more recent and largely limited to sequential techniques. Sun et al. [47] previously considered a one-dimensional (1D), two-equation, incompressible, and resistive MHD description of the geodynamo using an optimal interpolation DA scheme. Additionally, Fournier et al. [48] studied a similar 1D MHD model using a variational DA approach. Ren et al. [49] also considered the application of adjoint-based parameters and state estimation to a 1D Hartmann flow MHD system. Several DA studies pertaining to MHD have also been conducted for problems in two spatial dimensions (2D). Mendoza et al. [50] studied the application of the EnKF method to the ideal MHD equations for a 2D problem representative of the magnetospheric bow-shock of Earth. Synthetic data observed at various locations within the computational domain were used to constrain the simulations. While this previous study does not describe how the zero divergence condition for the magnetic field is enforced when applying the EnKF algorithm, Teixeira et al. [51] performed a subsequent 2D study using an equality constrained unscented Kalman filter method, allowing the divergence-free property of the magnetic field to be strictly enforced. Additionally, Biswas et al. [52] formulated a nudging DA scheme for the 2D incompressible and resistive MHD equations and Hudson et al. [53] carried out a related follow-up study that implemented and tested the DA scheme of the previous authors [52] using synthetic data. Merkin et al. [54] also examined sequential assimilation of low-altitude magnetic observations in the near-Earth environment with a linearized, 2D, steady-state approximation of the full MHD equations. Lastly, Lang and Owens [33] apply variational DA to a simplified model of the solar-wind based on the 2D inviscid Burgers equation [55]. The simplified single-equation solar-wind model was used due to the difficulty of obtaining an adjoint model for the full ideal MHD equations. The assimilation of both synthetic and in situ heliospheric observations were considered and, as mentioned above, Lang and Owens [33] demonstrated a strength of a variational approach applied to the inner boundary conditions for the simplified solar wind

model: the variational approach allowed observational data that are not directly upstream from Earth to improve near-Earth forecasts.

A few previous studies have also considered the application of DA methods to three-dimensional (3D) forms of the MHD equations. Li et al. [56] examined a 3D variant of the incompressible and resistive MHD system of the geodynamo to which a variational DA approach was applied. For the latter, the divergence-free conditions for the magnetic and velocity fields were handled explicitly in the definition of the cost function. More recently, Lang et al. [57] employed a local ensemble transform Kalman filter (LETKF) algorithm to demonstrate the first use of data assimilation with the ideal MHD equations in a three-dimensional setting. Lang et al. [57] utilized the previously mentioned ENLIL global MHD forecast model [13–15] and assimilated synthetic measurements of plasma density, temperature, and momentum at a single location in the heliosphere. Synthetic magnetic field observations were not included because the LETKF is not generally able to assimilate divergence-free fields, such as the magnetic field. Better agreement to near-Earth solar wind conditions were brought about by application of the DA procedure based on the upstream observations; however, the improvements diminished with time until the next set of observations were assimilated. In addition, Jivani et al. [58] recently conducted sensitivity analyses and uncertainty quantification for background solar wind properties using the aforementioned SWMF global MHD model. Although DA strategies were not considered, ensemble-based DA strategies can benefit from reduced sets of solution parameters as identified by Jivani et al. [58].

Finally, several authors have conducted research related to DA for space weather applications that were not based on global MHD descriptions. For example, Arge et al. [59] developed a framework for assimilating photospheric observations with a photospheric flux transport model for improved photospheric magnetograms. Hickmann et al. [60] subsequently advanced the work of Arge et al. [59] by incorporating an ensemble transform Kalman filter technique. Additionally, Innocenti et al. [61] applied a Kalman filtering technique to an empirical solar wind model for improved solar wind forecasts and Meadors et al. [62] used a particle filter method to determine the optimal radii of the source and interface surfaces of potential magnetic field models of the inner heliosphere that minimize discrepancies between the predictions of a Wang-Sheeley-Arge (WSA) empirical model of the solar wind [63,64] and solar wind observations.

2. Scope of current study

With the overarching motivation for improved space weather forecasting capabilities and preceding review in mind, this study considers the data assimilation of synthetic observational data to constrain solutions of 1D initial-value problems of the ideal MHD equations and performs a systematic and quantitative comparison of EnKF and 4DVar DA techniques. In particular, the DA methods are used is to reconstruct the correct flow-field given noisy sparse measurements following initiation with incorrect initial data for several problems including those with strongly non-linear behaviour and shocks. A standard upwind finite-volume spatial discretization procedure very representative of those used in the space weather prediction frameworks described in the introduction is used to obtain numerical solutions to the initial-value problems. Note that, to the author's knowledge, a direct comparison of EnKF and 4DVar approaches have not previously been considered for the ideal MHD equations and the application of 4DVar DA to ideal MHD is novel. For the latter, the formulation of an easily adaptable and implementable adjoint model for the ideal MHD equations with the upwind finite-volume scheme is also described. In addition, challenges associated with the assimilation of magnetic field data are also formally considered. It is noted that the "4D" in "4DVar" here formally refers to 3D physical space plus time. While the current study considers the application of a variational approach to unsteady solutions of the 1D

form of the ideal MHD equations, the technique is still referred to as 4DVar DA herein as seems to be the convention.

The organization of the remainder of the paper is as follows. Section 3 presents and summarizes the governing ideal MHD equations of interest. Section 4 then describes the upwind finite-volume spatial discretization and time-marching schemes used to construct numerical approximations to unsteady solutions of the ideal MHD equations on 1D domains. A probabilistic formalism of DA is then introduced in Section 5, which is subsequently used to describe both the EnKF and 4DVar DA algorithms considered herein. This is followed in Section 6 by a description and validation of the semi-analytical adjoint model for the discrete form of the 1D ideal MHD equations. In Section 7, multiple strategies are considered for mitigating non-zero magnetic field divergence errors brought about by the data assimilation. Section 8 reports on the comparisons of the EnKF and 4DVar schemes applied to several initial-value problems for the ideal MHD equations and, finally, Section 9 provides conclusions and suggestions for future research.

3. Ideal magnetohydrodynamics (MHD) equations

The equations of ideal MHD describe a perfectly electrically conducting inviscid plasma [65]. The mixture of ion and electron species are treated as a fully-ionized, quasi-neutral, single fluid having a Maxwellian particle distribution described by a single plasma temperature. Despite the many assumptions, the ideal MHD description remains a useful and self-consistent approximation for many astrophysical, space-physics, and engineering phenomena. In their non-dimensional and 1D form, the ideal MHD equations are a coupled hyperbolic system of non-linear partial differential equations that may be written as

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = \mathbf{0}, \tag{1}$$

where $\mathbf{U}$ is the conserved variable solution vector, and $\mathbf{F}$ is the system flux dyad. The conserved variable solution vector is given by

$$\mathbf{U} = [\rho, \rho\vec{u}, \vec{B}, E]^{\intercal}, \tag{2}$$

where ρ is the plasma density, $\vec{u} = [u_x, u_y, u_z]^{\intercal}$ is the velocity field, E is the total energy and $\vec{B} = [B_x, B_y, B_z]^{\intercal}$ is the magnetic field. Although only 1D problems are considered herein, all components of the velocity and magnetic field are accounted for here and allowed to vary along a single spatial dimension. Accordingly, the solution flux dyad is given by

$$\mathbf{F} = \begin{bmatrix} \rho\vec{u} \\ \rho\vec{u}\vec{u} - \vec{B}\vec{B} + (p + \frac{\vec{B}\cdot\vec{B}}{2})\vec{I} \\ \vec{B}\vec{u} - \vec{u}\vec{B} \\ \vec{u}(E + p + \frac{\vec{B}\cdot\vec{B}}{2}) - \vec{B}(\vec{u}\cdot\vec{B}) \end{bmatrix}, \tag{3}$$

where p is the plasma pressure and $\vec{I}$ is the identity dyad. A polytropic (thermally and calorically perfect) approximation is assumed for describing the thermodynamic behaviour of the plasma and thus the total energy and plasma pressure are related according to

$$E = \frac{p}{\gamma - 1} + \frac{\rho\vec{u}\cdot\vec{u}}{2} + \frac{\vec{B}\cdot\vec{B}}{2}, \tag{4}$$

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. The solutions of Eq. (1) are also subject to the divergence-free constraint on the magnetic field, $\vec{\nabla}\cdot\vec{B} = 0$. In the 1D setting of interest here, this reduces to $\partial B_x/\partial x = 0$ or $B_x = $ constant.

4. Upwind finite-volume solution procedure without $\vec{\nabla} \cdot \vec{B} = 0$ treatment

A Godunov-type, upwind, finite-volume scheme [66] is considered for the numerical solution of the hyperbolic system of Eq. (1). 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, and are a key component of many of the space weather forecasting tools based on global MHD modelling. The latter includes the SWMF [4–12], EUHFORIA [16,17], and framework of Narechania *at al.* [18] discussed previously. For example, the upwind scheme of Powell et al. [4] is the basis for the solution of the MHD equations in the SWMF simulation code and both second- and high-order, upwind, finite-volume variants [67–73] can be used within the simulation tool of Narechania *at al.* [18].

In this study, a limited, second-order, MUSCL-type, upwind, finite-volume scheme [74,75] is considered as the representative scheme for the numerical solution of Eq. (1) on uniform 1D mesh. It is very similar to the multidimensional second-order scheme with limiting used in the discretization of the ideal MHD equations by Ivan et al. [67]. In the 1D case for which the $\vec{\nabla} \cdot \vec{B} = 0$ constraint is explicitly strictly enforced based on the given initial data, application of this upwind spatial discretization procedure to the governing equations for the ith computational cell of the 1D grid results in the following semi-discrete form of the MHD equations given by

$$\frac{d\mathbf{U}_i}{dt} = -\frac{1}{\Delta x}(\mathcal{F}_{i+1/2} - \mathcal{F}_{i-1/2}) \equiv \mathbf{R}_i(\mathbf{U}), \tag{5}$$

where $\mathcal{F}_{i\pm1/2}$ are the upwind values of the solution fluxes evaluated at the $i \pm 1/2$ cell boundary or interface, $\mathbf{U}_i$ is the cell-averaged value of the conserved solution vector within cell i, and $\mathbf{R}_i$ is defined to be the steady residual for cell i. The numerical fluxes are evaluated via the solution of Riemann problems defined in terms of the primitive solution values, $\mathbf{W}_L$ and $\mathbf{W}_R$, to the left and right of the interface, respectively, with $\mathbf{W} = [\rho, \vec{u}, \vec{B}, p]^\mathsf{T}$, such that $\mathcal{F}_{i\pm1/2} = \mathcal{F}_{i\pm1/2}(\mathbf{W}_{L,i\pm1/2}, \mathbf{W}_{R,i\pm1/2})$. The approximate Riemann solver of Powell [4,76] is used herein to evaluate the numerical fluxes and second-order spatial accuracy is achieved via piece-wise limited linear reconstruction of the primitive solution vector, $\mathbf{W}$, at the left and right cell interfaces according to

$$\mathbf{W}_{L,i+1/2} = \mathbf{W}_{i-1/2\pm1/2} + \phi_{i-1/2\pm1/2} \circ \frac{\partial\mathbf{W}_{i-1/2\pm1/2}}{\partial x}\frac{\Delta x}{2}, \tag{6}$$

$$\mathbf{W}_{R,i+1/2} = \mathbf{W}_{i+1/2\pm1/2} - \phi_{i+1/2\pm1/2} \circ \frac{\partial\mathbf{W}_{i+1/2\pm1/2}}{\partial x}\frac{\Delta x}{2}, \tag{7}$$

where $\circ$ here denotes an element-wise product, and ϕ is the Van Leer slope limiter [74]. The solution gradient, $\partial\mathbf{W}/\partial x$, is specified by the method of least squares and is given by

$$\frac{\partial\mathbf{W}_i}{\partial x} = \frac{\Delta\mathbf{W}_i\Delta x_i - \nabla\mathbf{W}_i\nabla x_i}{(\Delta x_i)^2 + (\nabla x_i)^2}, \tag{8}$$

where Δx_i and ∇x_i denote forward and backwards differences respectively.

For the unsteady initial-value problems of interest here, the semi-discrete form of Eq. (1) is solved by adopting a method-of-lines approach and applying a Runge–Kutta time-marching scheme. Second-order temporal accuracy is achieved here by using a strong stability preserving (SSP), second-order, Runge–Kutta, time-marching scheme [77,78] given by

$$\widetilde{\mathbf{U}}^{n+1} = \mathbf{U}^n + \Delta t\mathbf{R}(\mathbf{U}^n), \tag{9}$$

$$\mathbf{U}^{n+1} = \frac{1}{2}(\mathbf{U}^n + \widetilde{\mathbf{U}}^{n+1} + \Delta t\mathbf{R}(\widetilde{\mathbf{U}}^{n+1})), \tag{10}$$

where $\mathbf{U}^n$ stores the cell-averaged conserved solution vectors for the entire grid at the nth 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 criteria [79]. Similar to the steady residual, $\mathbf{R}$, we define the unsteady residual, $\mathbf{R}^*$, by rearranging Eqs. (9) and (10), yielding

$$\mathbf{R}^* \equiv \frac{\mathbf{U}^{n+1} - \mathbf{U}^n}{\Delta t} - \frac{\mathbf{R}(\mathbf{U}^n) + \mathbf{R}(\widetilde{\mathbf{U}}^{n+1})}{2} = 0. \tag{11}$$

5. Data assimilation (DA) without $\vec{\nabla} \cdot \vec{B} = 0$ treatment

Data assimilation of sparse measured data associated with the time evolution of the solution for a general physical system is now considered. Given separate estimates of a physical system, namely a dynamical model, $\mathcal{M}$, and the measured data, $\mathbf{d}$, an optimal trajectory of the system solution or state, $\mathbf{U}$, is sought that utilizes the model, data, and the associated uncertainties associated with both. The focus of this study is on fully discrete dynamical models of the form

$$\mathbf{U}^{n+1} = \mathcal{M}(\mathbf{U}^n) + \boldsymbol{\eta}^n, \quad 0 \le n \le N - 1, \tag{12}$$

with initial conditions

$$\mathbf{U}^0 = \mathbf{U}^0_{\text{truth}} + \boldsymbol{\zeta}, \tag{13}$$

and observations

$$\mathbf{d}^n = \mathcal{H}(\mathbf{U}^n_{\text{truth}}) + \boldsymbol{\xi}^n, \quad 1 \le n \le N, \tag{14}$$

where $\mathbf{U}^{n+1}$ and $\mathbf{U}^n$ represents the solution or state at time levels $n + 1$ and n, respectively, N is the total number of discrete time levels of interest, and $\boldsymbol{\eta}^n$, $\boldsymbol{\zeta}$, and $\boldsymbol{\xi}^n$ are stochastic error terms with known distributions corresponding to the model, initial condition, and measurements, respectively. The observation operator, $\mathcal{H} : \mathbb{R}^{N_{\mathbf{U}}} \to \mathbb{R}^{N_{\mathbf{d}}}$, maps the state space to the observation space, and $\mathbf{U}^n_{\text{truth}}$ is a discretization of the true state at the nth time step. Errors in the model, $\boldsymbol{\eta}^n$, may be due to discretization errors, missing physics, or inaccurate parameters for example. Similarly, errors in the initial condition, $\boldsymbol{\zeta}$, arise from incomplete or unreliable knowledge of the initial state. Additionally, both instrument noise and the restricted representation of the system state on a computational grid contribute to observational errors, $\boldsymbol{\xi}^n$. In the present study, the mathematical model errors, $\boldsymbol{\eta}^n$, are not considered. The errors in the initial condition and observations are assumed to be Gaussian with zero mean and known covariance matrices, $\boldsymbol{\Sigma}^0_{\mathbf{U}}$ and $\boldsymbol{\Sigma}_{\mathbf{d}}$, respectively, such that

$$\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}^0_{\mathbf{U}}), \tag{15}$$

$$\boldsymbol{\xi}^n \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}^n_{\mathbf{d}}). \tag{16}$$

In addition, observations are linearly related to the model state and thus the observation operator, $\mathcal{H}$, may be replaced by an observation matrix, $\mathbf{H}$.

The system state and observations are treated as random variables with the *a priori* probability density $f(\mathbf{U}^0, \dots, \mathbf{U}^N)$, and likelihood $f(\mathbf{d}^1, \dots, \mathbf{d}^N | \mathbf{U}^0, \dots, \mathbf{U}^N)$. Following Van Leeuwen and Evensen [80], the posterior distribution of the system trajectory, given the data, can be expressed through Bayes' theorem as

$$f(\mathbf{U}^0, \dots, \mathbf{U}^N | \mathbf{d}^1, \dots, \mathbf{d}^N) \propto f(\mathbf{U}^0, \dots, \mathbf{U}^N)f(\mathbf{d}^1, \dots, \mathbf{d}^N | \mathbf{U}^0, \dots, \mathbf{U}^N). \tag{17}$$

The proportionality constant is omitted as it is simply a normalizing constant and does not affect the Bayesian estimation problem. Although the computation of Eq. (17) is not feasible, one is generally interested in the related moments rather than the probability density function itself. In fact, the optimal trajectory can be determined by finding the value of $\mathbf{U}^0, \dots, \mathbf{U}^N$ that maximizes $f(\mathbf{U}^0, \dots, \mathbf{U}^N | \mathbf{d}^1, \dots, \mathbf{d}^N)$, known as the maximum a posteriori estimate. In addition, the minimum-variance estimate is equal to the maximum a posteriori estimate when the variables follow a Gaussian distribution.

Data assimilation methods can readily be interpreted by examining how Eq. (17) above is treated. In the following subsections, the EnKF and 4DVar are introduced by first making modifications to Eq. (17).

5.1. Ensemble Kalman filter (EnKF) technique

Adopting the same probabilistic formalism used to derive Eq. (17), one may obtain a general Bayesian filter [81] given by

$$f(\mathbf{U}^n | \mathbf{d}^1, \dots, \mathbf{d}^n) \propto f(\mathbf{U}^n | \mathbf{d}^1, \dots, \mathbf{d}^{n-1})f(\mathbf{d}^n | \mathbf{U}^n). \tag{18}$$

Three assumptions are needed in the derivation of this general filter, namely (i) the model evolution is a first-order Markov process, i.e.,

$$f(\mathbf{U}^n|\mathbf{U}^0,\dots,\mathbf{U}^{n-1}) = f(\mathbf{U}^n|\mathbf{U}^{n-1}); \tag{19}$$

(ii) the data is temporally independent, i.e.,

$$f(\mathbf{d}^1,\dots,\mathbf{d}^N|\mathbf{U}^0,\dots,\mathbf{U}^N) = \prod_{n=1}^N f(\mathbf{d}^n|\mathbf{U}^0,\dots,\mathbf{U}^N); \tag{20}$$

and (iii) data at time step, n, only depends on the state at the same time step, i.e.,

$$f(\mathbf{d}^n|\mathbf{U}^0,\dots,\mathbf{U}^N) = f(\mathbf{d}^n|\mathbf{U}^n). \tag{21}$$

Compared with Eq. (17), the Bayesian filter seeks the optimal value of the state at time, t^n, from knowledge of the prior data assimilation and time integration steps, as well as the data at time, t^n. This process is applied sequentially.

When the error statistics are Gaussian and the dynamical model is linear, the original KF method [24] provides an estimate, $\hat{\mathbf{U}}^n$, that maximizes the probability density of Eq. (18). An important limitation of the KF is that it cannot be employed with non-linear dynamics, even suboptimally. The EnKF method, on the other hand, uses an ensemble representation of the prior probability densities which allows for the non-linear integration of each ensemble member. In the limit of an infinite ensemble size and linear dynamics, the EnKF approach is equivalent to the KF method. However, the posterior probability density does not remain Gaussian when the model is non-linear, even with Gaussian priors. Consequently, the minimum-variance estimate produced by the EnKF is not equivalent to the maximum a posteriori estimate. Nonetheless, the EnKF remains a suitable and effective algorithm for many problems.

The EnKF can be divided into two steps: the forecast and analysis steps. In the former, the ensemble is propagated forward in time by applying Eq. (12) to each of the S ensemble members. In the present study, the dynamical model, $\mathcal{M}$, represents the ideal MHD solution scheme of Eq. (10), and each ensemble member, $\mathbf{U}^{n,s}$, represents a possible realization of the conserved quantities in the MHD flow. The analysis step involves the addition of a linear correction term to each ensemble member, i.e.,

$$\hat{\mathbf{U}}^{n,s} = \mathbf{U}^{n,s} + \mathbf{K}^n(\mathbf{d}^{n,s} - \mathbf{H}^n\mathbf{U}^{n,s}), \quad \forall s \in \{1,\dots,S\}, \tag{22}$$

where an ensemble of observations is constructed according to

$$\mathbf{d}^{n,s} = \mathbf{d}^n + \xi^{n,s}, \tag{23}$$

and the Kalman gain, $\mathbf{K}^n$, is given by

$$\mathbf{K}^n = \Sigma_{\mathbf{U}}^n \mathbf{H}^{n\mathsf{T}}(\mathbf{H}^n \Sigma_{\mathbf{U}}^n \mathbf{H}^{n\mathsf{T}} + \Sigma_{\mathbf{d}}^n)^{-1}. \tag{24}$$

As a reminder, the matrices, $\Sigma_{\mathbf{U}}^n$ and $\Sigma_{\mathbf{d}}^n$, are the covariance matrices of the model and data respectively. While $\Sigma_{\mathbf{d}}^n$ is known *apriori*, $\Sigma_{\mathbf{U}}^n$ is estimated by the sample covariance

$$\Sigma_{\mathbf{U}}^n = \frac{\mathbf{A}^n \mathbf{A}^{n\mathsf{T}}}{S-1}, \tag{25}$$

with

$$\overline{\mathbf{U}^n} = \frac{1}{S}\sum_{s=1}^S \mathbf{U}^{n,s}, \qquad \mathbf{A}^n = [\mathbf{U}^{n,1} - \overline{\mathbf{U}^n},\dots,\mathbf{U}^{n,S} - \overline{\mathbf{U}^n}]. \tag{26}$$

Due to the small number of ensemble members, the sample covariance often produces spurious correlations between distant grid-points. To reduce or dampen these erroneous correlations, a so-called localization matrix, $\mathbf{L}$, is used to multiply element-wise the sample covariance to arrive at a localized covariance matrix given by

$$(\Sigma_{\mathbf{U}}^n)_{\text{localized}} = \mathbf{L} \circ \Sigma_{\mathbf{U}}^n. \tag{27}$$

The localization matrix, $\mathbf{L}$, is constructed such that variances remain unchanged, and covariances approach zero with increasing distance. In

the present work, the second-order autoregressive distribution is used to specify $\mathbf{L}$. Accordingly, the localization matrix is given by

$$\mathbf{L}(i,j) = \left(1 + \frac{|i-j|\Delta x}{D_L}\right) \exp\left(\frac{-|i-j|\Delta x}{D_L}\right), \tag{28}$$

where D_L is the localization length scale.

5.2. Four-dimensional variational (4DVar) technique

The so-called *strong-constraint* 4DVar method seeks optimal initial data, $\mathbf{U}^0$, for cases in which the model errors, η, are ignored. Under these conditions, Eq. (17) reduces to

$$f(\mathbf{U}^0|\mathbf{d}^1,\dots,\mathbf{d}^N) \propto f(\mathbf{U}^0)f(\mathbf{d}^1,\dots,\mathbf{d}^N|\mathbf{U}^0) \propto \exp(-J(\mathbf{U}^0)). \tag{29}$$

The 4DVar approach thus calculates the maximum a posteriori estimate of the initial condition. In the case of a linear model, this optimal choice for the initial data provides the most-likely estimate of $\mathbf{U}$ as time evolves. For a non-linear model, the optimal initial solution does not ensure the most-likely estimates for $\mathbf{U}$ over the time window of interest and, for this reason, like the EnKF approach the 4DVar method does not produce a strictly optimal non-linear model trajectory. Despite this limitation, the 4DVar technique remains a useful tool for practical data assimilation.

For the 4DVar approach, it can be shown that the posterior probability density is maximized by minimizing the cost-function, $J(\mathbf{U}^0)$, given by

$$J(\mathbf{U}^0) = \frac{1}{2}(\mathbf{U}^0 - \mathbf{U}^b)^\mathsf{T}(\Sigma_{\mathbf{U}}^0)^{-1}(\mathbf{U}^0 - \mathbf{U}^b) + \frac{1}{2}\sum_{n=1}^N (\mathbf{H}^n\mathbf{U}^n - \mathbf{d}^n)^\mathsf{T}(\Sigma_{\mathbf{d}}^n)^{-1}(\mathbf{H}^n\mathbf{U}^n - \mathbf{d}^n), \tag{30}$$

while simultaneously constrained by the model, $\mathbf{U}^{n+1} = \mathcal{M}(\mathbf{U}^n)$. Here, $\mathbf{U}^b$ denotes the background (*a priori*) estimate of the initial condition. One may interpret J as a measure of two competing quantities. The first term measures the squared difference between the posterior and prior initial conditions, weighed by the precision of the model. The remaining terms measure the squared difference between the model and the data, weighed by the precision of the data. In the context of the finite-volume solution scheme of Section 4, the 4DVar estimate can be expressed as the solution of the following minimization problem

$$\underset{\mathbf{U}^0}{\text{minimize}} \quad J(\mathbf{U}^0), \tag{31a}$$

$$\text{subject to} \quad \mathbf{R}^* = \mathbf{0}. \tag{31b}$$

Rather than solving the constrained problem directly, Lagrange multipliers, often termed adjoint variables, are introduced to define the Lagrangian, $\mathcal{L}$, given by

$$\mathcal{L}(\mathbf{U}^0,\dots,\mathbf{U}^{N+1},\lambda^1,\dots,\lambda^{N+1}) = J + \sum_{n=0}^N (\lambda^{n+1})^\mathsf{T}\mathbf{R}^*(\mathbf{U}^{n+1},\mathbf{U}^n), \tag{32}$$

as a function of all solution variables, $\mathbf{U}^n$, and all adjoint variables, λ^n. Solutions to the constrained problem may be obtained by finding the stationary points of the Lagrangian. For this, the derivatives of $\mathcal{L}$ with respect to each of the $N+2$ solution variables and $N+1$ adjoint variables must vanish. Before enforcing these conditions on the derivatives of $\mathcal{L}$, it is instructive to note that the unsteady residual may be written as

$$\mathbf{R}^*(\mathbf{U}^{n+1},\mathbf{U}^n) = \mathbf{U}^{n+1} - \mathcal{M}(\mathbf{U}^n), \tag{33}$$

where $\mathcal{M}(\mathbf{U}^n)$ is given by the right hand of Eq. (10). In this form, Jacobians of $\mathbf{R}^*$ can trivially be written as

$$\frac{\partial \mathbf{R}^*(\mathbf{U}^{m+1},\mathbf{U}^m)}{\partial \mathbf{U}^n} = \mathbf{0}, \quad \forall m \notin \{n, n-1\}, \tag{34}$$

$$\frac{\partial \mathbf{R}^*(\mathbf{U}^{n+1},\mathbf{U}^n)}{\partial \mathbf{U}^n} = -\frac{\partial \mathcal{M}(\mathbf{U}^n)}{\partial \mathbf{U}^n}, \qquad \frac{\partial \mathbf{R}^*(\mathbf{U}^n,\mathbf{U}^{n-1})}{\partial \mathbf{U}^n} = \mathbf{I}. \tag{35}$$

Now, derivatives of $\mathcal{L}$ may be written as

$$\frac{\partial \mathcal{L}}{\partial \lambda^n} = \mathbf{R}^*(\mathbf{U}^{n+1}, \mathbf{U}^n), \qquad \forall n \in \{1, \ldots, N\}, \tag{36}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{U}^{N+1}} = \lambda^{N+1}, \tag{37}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{U}^n} = \mathbf{H}^{n\mathsf{T}}(\Sigma_\mathbf{d}^n)^{-1}(\mathbf{H}^n\mathbf{U}^n - \mathbf{d}^n) - \left[\frac{\partial \mathcal{M}}{\partial \mathbf{U}^n}\right]^\mathsf{T} \lambda^{n+1} + \lambda^n, \qquad \forall n \in \{1, \ldots, N\}, \tag{38}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{U}^0} = (\Sigma_\mathbf{U}^0)^{-1}(\mathbf{U}^0 - \mathbf{U}^b) - \left[\frac{\partial \mathcal{M}}{\partial \mathbf{U}^0}\right]^\mathsf{T} \lambda^1$$
$$= (\Sigma_\mathbf{U}^0)^{-1}(\mathbf{U}^0 - \mathbf{U}^b) - \lambda^0, \tag{39}$$

where the extra variable, λ^0, has been introduced for convenience. Setting Eqs. (36)–(39) all to zero yields the coupled system of Euler–Lagrange equations

$$\mathbf{R}^*(\mathbf{U}^n, \mathbf{U}^{n-1}) = \mathbf{0}, \qquad \forall n \in \{1, \ldots, N\}, \tag{40}$$

$$\lambda^{N+1} = \mathbf{0}, \tag{41}$$

$$\lambda^n = \left[\frac{\partial \mathcal{M}}{\partial \mathbf{U}^n}\right]^\mathsf{T} \lambda^{n+1} - \mathbf{H}^{n\mathsf{T}}(\Sigma_\mathbf{d}^n)^{-1}(\mathbf{H}^n\mathbf{U}^n - \mathbf{d}^n), \qquad \forall n \in \{1, \ldots, N\}, \tag{42}$$

$$\lambda^0 = (\Sigma_\mathbf{U})^{-1}(\mathbf{U}^0 - \mathbf{U}^b). \tag{43}$$

Eqs. (41)–(42) are known as the adjoint equations and form a two-point boundary-value problem. Noticeably, the fully discrete ideal MHD equations are recovered via Eq. (40). In practice, rather than solving the two-point boundary-value problem directly, the cost-function gradient, $\partial J/\partial \mathbf{U}^0$, is calculated by use of Eq. (39) given that $\partial \mathcal{L}/\partial \mathbf{U}^n = \mathbf{0}$ and $\partial \mathcal{L}/\partial \lambda^n = \mathbf{0}$ are satisfied exactly for all time steps. These conditions are ensured by the integration of the adjoint equations and fully discrete ideal MHD equations, allowing one to write

$$\frac{\partial J}{\partial \mathbf{U}^0} = \frac{\partial \mathcal{L}}{\partial \mathbf{U}^0}\bigg|_{\frac{\partial \mathcal{L}}{\partial \mathbf{U}^n} = \frac{\partial \mathcal{L}}{\partial \lambda^n} = \mathbf{0}} = (\Sigma_\mathbf{U}^0)^{-1}(\mathbf{U}^0 - \mathbf{U}^b) - \lambda^0. \tag{44}$$

With knowledge of the gradient, $\partial J/\partial \mathbf{U}^0$, the cost-function can be iteratively minimized by a gradient based optimization algorithm until Eq. (43) is satisfied. The limited-memory variant of the Broyden–Fletcher–Goldfarb– Shanno (L-BFGS) algorithm [82], a limited-memory quasi-Newton method, is used in the present study to drive the minimization process. Faster convergence is achieved by preconditioning the problem as outlined in Appendix A.

5.3. Summary of data assimilation (DA) algorithms

Summaries of both the EnKF and 4DVar DA algorithms formulated and described above and applied herein to the ideal MHD initial-value problems of interest are given in Table 1.

6. Formulation of discrete adjoint model without $\vec{\nabla} \cdot \vec{B} = 0$ treatment

A common difficulty with the implementation of any 4DVar approach is the formulation of the transpose of the linearized model, $[\partial \mathcal{M}/\partial \mathbf{U}^n]^\mathsf{T}$, appearing in the adjoint equations of Eq. (42). This term, often referred to as the discrete adjoint model, can be both involved and challenging to evaluate. In particular, for discrete models of the physical system as considered here, the derivatives of the functions defined in every line of computer code must be evaluated and transposed. Various strategies are possible for the evaluation of the discrete adjoint model, including direct (i.e., by hand) analytical differentiation, the use of finite-difference approximations, the complex-step derivative approximation, and the use of various automatic differentiation libraries [83–85]. It is important to note that adjoint-based techniques are used extensively in many other optimal control applications ranging from aerodynamic shape and multidisciplinary design optimization to output-based error estimation for directing mesh adaptation. See, for

example, the previous research by Pironneau [86], Jameson [87], Nemec and Zingg [88,89], Truong et al. [90], and Hicken and Zingg [91] related to design optimization and the previous studies of Becker and Rannacher [92,93], Becker et al. [94], Heuveline and Rannacher [95], Venditti and Darmofal [96–98], Nemec and Aftosmis [99,100], and Ceze and Fidkowski [101]. For conducting gaseous flows, Marta and Alonso [102–104] developed discrete adjoint solution methods for the ideal MHD and low-magnetic Reynolds-number regimes in order to perform sensitivity analysis and optimize hypersonic vehicle designs while accounting for magnetic effects on locally ionized flows. More recently, Narechania [105] developed a discrete adjoint method for steady ideal MHD flow for use in the first application of output-based error estimation with anisotropic, block-based, adaptive mesh refinement (AMR) to plasma flows.

A hybrid approach is adopted here for the evaluation of the discrete adjoint model associated with the 1D form of the ideal MHD equations. In particular, the finite-volume solution scheme of Section 4 is differentiated and transposed analytically with the exception of the numerical flux function, $\mathcal{F}$. In the case of the latter, derivatives of $\mathcal{F}$ are then computed using the ADEPT automatic differentiation library [85]. With this hybrid approach, the adjoint model is easily adaptable to variations in the scheme and even in the governing equations. In addition, by presenting and evaluating the discrete adjoint model in proposed form, it is readily replicated in possible follow-on studies.

To begin the formulation of the discrete adjoint model, the focus will first be on the temporal discretization. Differentiation of the time marching scheme is independent of the spatial discretization strategy, allowing for modularity in the schemes of choice. Differentiating the second-order SSP Runge–Kutta scheme of Eqs. (9) and (10) yields

$$\widetilde{\lambda}^n = \lambda^{n+1} + \Delta t \left[\frac{\partial \mathbf{R}(\widetilde{\mathbf{U}}^{n+1})}{\partial \widetilde{\mathbf{U}}^{n+1}}\right]^\mathsf{T} \lambda^{n+1}, \tag{45}$$

$$\left[\frac{\partial \mathcal{M}}{\partial \mathbf{U}^n}\right]^\mathsf{T} \lambda^{n+1} = \frac{1}{2}\left(\lambda^{n+1} + \widetilde{\lambda}^n + \Delta t \left[\frac{\partial \mathbf{R}(\mathbf{U}^n)}{\partial \mathbf{U}^n}\right]^\mathsf{T} \widetilde{\lambda}^n\right), \tag{46}$$

which takes a familiar predictor–corrector structure. The derivation of Eqs. (45) and (46) is provided in Appendix B. In the above, the adjoint model is now written solely in terms of the Jacobian of the steady residual, $\partial \mathbf{R}/\partial \mathbf{U}$. To evaluate this Jacobian, three vectors are introduced. The first vector, $\mathcal{F} = [\mathcal{F}_{1/2}, \ldots, \mathcal{F}_{i-1/2}, \ldots, \mathcal{F}_{M+1/2}]^\mathsf{T}$, contains the numerical fluxes at the $M + 1$ cell interfaces. The elements of the second and third vectors, $\mathbf{W}_L = [\mathbf{W}_{L,1/2}, \ldots, \mathbf{W}_{L,i-1/2}, \ldots, \mathbf{W}_{L,M+1/2}]^\mathsf{T}$, and $\mathbf{W}_R = [\mathbf{W}_{R,1/2}, \ldots, \mathbf{W}_{R,i-1/2}, \ldots, \mathbf{W}_{R,M+1/2}]^\mathsf{T}$, are the reconstructed primitive solution variables associated with the left and right cell interfaces, respectively. The steady residual may now be written as a function of $\mathcal{F}$, i.e.,

$$\mathbf{R}^n = \mathbf{R}(\mathcal{F}(\mathbf{W}_L^n, \mathbf{W}_R^n)), \tag{47}$$

where $\mathcal{F}$ is in turn dependent on the conserved solution variables via

$$\mathbf{W}_L^n = \mathbf{W}_L(\mathbf{W}(\mathbf{U}^n)), \qquad \mathbf{W}_R^n = \mathbf{W}_R(\mathbf{W}(\mathbf{U}^n)). \tag{48}$$

With these preliminaries, and after successive application of the chain rule, one may write

$$\frac{\partial \mathbf{R}^n}{\partial \mathbf{U}^n} = \frac{\partial \mathbf{R}^n}{\partial \mathcal{F}^n}\left(\frac{\partial \mathcal{F}^n}{\partial \mathbf{W}_L^n}\frac{\partial \mathbf{W}_L^n}{\partial \mathbf{W}^n} + \frac{\partial \mathcal{F}^n}{\partial \mathbf{W}_R^n}\frac{\partial \mathbf{W}_R^n}{\partial \mathbf{W}^n}\right)\frac{\partial \mathbf{W}^n}{\partial \mathbf{U}^n}. \tag{49}$$

With the exception of $\partial \mathcal{F}^n/\partial \mathbf{W}_L^n$ and $\partial \mathcal{F}^n/\partial \mathbf{W}_R^n$, each of the Jacobian matrices are evaluated analytically, and the relevant expressions are given in Appendix B. Taking the transpose of Eq. (49), and noting that $\partial \mathcal{F}^n/\partial \mathbf{W}_L^n$ and $\partial \mathcal{F}^n/\partial \mathbf{W}_R^n$ are block diagonal quantities, the Jacobian-transpose-vector products appearing in Eqs. (45) and (46) may be evaluated directly without the full storage of these large matrices. Instead, the ith element of $\hat{\lambda} = [\partial \mathbf{R}/\partial \mathbf{U}]^\mathsf{T}\lambda$ is computed according to

Table 1
Comparison of data assimilation algorithms studied herein.

Algorithm 1 4DVar Method	**Algorithm 2** EnKF Method
1: **input:** $\mathbf{U}^b$, $\mathbf{d}$	1: **input:** $\mathbf{U}^b$, $\mathbf{d}$
2: $\mathbf{U}^0 = \mathbf{U}^b$	2: $\overline{\mathbf{U}^n} = \mathbf{U}^b$
3: **repeat**	3: **for** $s \in \{1, \cdots, S\}$ **do**
4: **for** $n \in \{0, \cdots, N-1\}$ **do**	4: $\mathbf{U}^{0,s} = \overline{\mathbf{U}^n} + \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{\mathbf{U}}^n)$
5: $\mathbf{U}^{n+1} = \mathcal{M}(\mathbf{U}^n)$	5: **end for**
6: **end for**	6: **for** $n \in \{1, \cdots, N\}$ **do**
7: $\boldsymbol{\lambda}^{N+1} = \mathbf{0}$	7: **for** $s \in \{1, \cdots, S\}$ **do**
8: **for** $n \in \{N, \cdots, 1\}$ **do**	8: $\mathbf{U}^{n,s} = \mathcal{M}(\mathbf{U}^{n-1,s})$
9: $\boldsymbol{\psi}^n = \mathbf{H}^{n\intercal}(\mathbf{\Sigma}_{\mathbf{d}}^n)^{-1}(\mathbf{H}^n\mathbf{U}^n - \mathbf{d}^n)$	9: **end for**
10: $\boldsymbol{\lambda}^n = [\frac{\partial \mathcal{M}}{\partial \mathbf{U}^n}]^{\intercal}\boldsymbol{\lambda}^{n+1} - \boldsymbol{\psi}^n$	10: $\overline{\mathbf{U}^n} = \frac{1}{S}\sum_{s=1}^{S}\mathbf{U}^{n,s}$
11: **end for**	11: $\mathbf{\Sigma}_{\mathbf{U}}^n = \mathbf{L} \circ \frac{\mathbf{A}^n \mathbf{A}^{n\intercal}}{S-1}$
12: $\boldsymbol{\lambda}^0 = [\frac{\partial \mathcal{M}}{\partial \mathbf{U}^n}]^{\intercal}\boldsymbol{\lambda}^1$	12: $\mathbf{K}^n = \mathbf{\Sigma}_{\mathbf{U}}^n \mathbf{H}^{n\intercal}(\mathbf{H}^n\mathbf{\Sigma}_{\mathbf{U}}^n\mathbf{H}^{n\intercal} + \mathbf{\Sigma}_{\mathbf{d}}^n)^{-1}$
13: $\frac{\partial J}{\partial \mathbf{U}^0} = (\mathbf{\Sigma}_{\mathbf{U}}^0)^{-1}(\mathbf{U}^0 - \mathbf{U}^b) - \boldsymbol{\lambda}^0$	13: **for** $s \in \{1, \cdots, S\}$ **do**
14: $\Delta\mathbf{U}^0 = \text{L-BFGS}(\frac{\partial J}{\partial \mathbf{U}^0})$	14: $\mathbf{d}^{n,s} = \mathbf{d}^n + \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{\mathbf{d}}^n)$
15: $\mathbf{U}^0 = \mathbf{U}^0 + \Delta\mathbf{U}^0$	15: $\mathbf{U}^{n,s} = \mathbf{U}^{n,s} + \mathbf{K}^n(\mathbf{d}^{n,s} - \mathbf{H}^n\mathbf{U}^{n,s})$
16: **until** converged	16: **end for**
	17: **end for**

$$\lambda_i = \frac{1}{\Delta x}\left[\frac{\partial \mathbf{W}_i}{\partial \mathbf{U}_i}\right]^{\intercal}\sum_{j=-1}^{1}\left[\frac{\partial \mathbf{W}_{L,i+j+1/2}}{\partial \mathbf{W}_i}\right]^{\intercal}\left[\frac{\partial \mathcal{F}_{i+j+1/2}}{\partial \mathbf{W}_{L,i+j+1/2}}\right]^{\intercal}\Delta\lambda_{i+j}$$
$$+ \left[\frac{\partial \mathbf{W}_{R,i+j-1/2}}{\partial \mathbf{W}_i}\right]^{\intercal}\left[\frac{\partial \mathcal{F}_{i+j-1/2}}{\partial \mathbf{W}_{R,i+j-1/2}}\right]^{\intercal}\nabla\lambda_{i+j},$$

$$(50)$$

where $\Delta\lambda_{i+j}$ and $\nabla\lambda_{i+j}$ denote the forward and backwards differences of λ_{i+j}. The Jacobian-transpose-vector products involving the numerical fluxes, $\mathcal{F}$, appearing in Eq. (50) are evaluated via ADEPT's reverse-mode automatic differentiation [85]. With the above defined, Eq. (42), together with Eqs. (45), (46), and (50), can then be used to integrate the discrete adjoint equations backwards in time thereby obtaining the adjoint solution at each discrete time step of a numerical simulation.

6.1. Validation of discrete adjoint model

To verify the correctness of discrete adjoint model formulation, a so-called gradient test was performed here. As described by Navon et al. [106], the accuracy of the gradient computed by the adjoint model can be assessed by considering a Taylor-expansion of the cost function given by

$$J(\mathbf{U}^0 + \alpha\delta\mathbf{U}) = J(\mathbf{U}^0) + \alpha\delta\mathbf{U}^{\intercal}\frac{\partial J}{\partial \mathbf{U}^0} + \mathcal{O}(\alpha^2),$$

$$(51)$$

where α is a small scalar, and $\delta\mathbf{U}$ is a unit vector. Re-writing the above expression, one may define the function, $\beta(\alpha)$, in terms of α as follows

$$\beta(\alpha) = \left|\frac{J(\mathbf{U}^0 + \alpha\delta\mathbf{U}) - J(\mathbf{U}^0)}{\alpha\delta\mathbf{U}^{\intercal}\frac{\partial J}{\partial \mathbf{U}^0}} - 1\right| = \mathcal{O}(\alpha),$$

$$(52)$$

which then quantifies the error in the gradient, $\partial J/\partial\mathbf{U}^0$. If the discrete adjoint model is formulated correctly, β should be small for a small α and converge to 0 with an error of order one. Choosing $\delta\mathbf{U} = [\partial J/\partial\mathbf{U}^0]/\|\partial J/\partial\mathbf{U}^0\|$, the computed value of β is depicted in Fig. 1 for the present adjoint model as a function of α. The error is shown to be small and converges with order one, thus validating the adjoint model formulation adopted herein. Note that, for very small values of α (i.e., $\alpha < 10^{-10}$), the value of β can be observed to increase again as finite-precision, numerical, round-off errors dominate and negate the accuracy of the validation strategy in these cases.
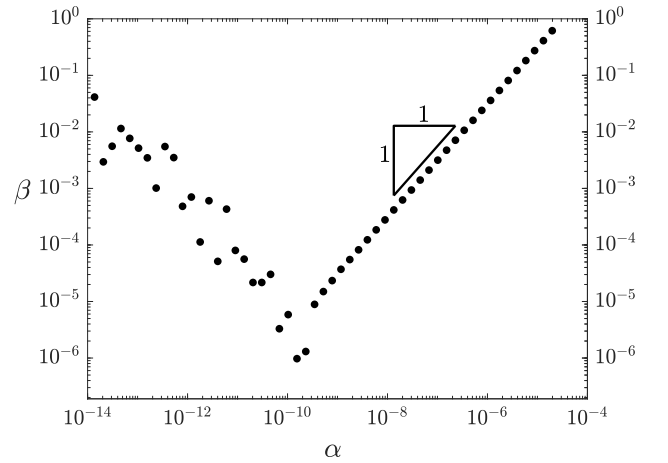
Fig. 1. Verification of the discrete adjoint model via the gradient test approach. β quantifies the error in the adjoint-based gradient evaluation given the step size, α.

7. Strategies for enforcement of $\vec{\nabla} \cdot \vec{B} = 0$ constraint with data assimilation

In space plasma physics as in many plasma applications, the magnetic field plays a crucial role. Its assimilation is therefore paramount for successful space plasma DA frameworks. One of the primary challenges in MHD-based simulations for space weather forecasting is the enforcement of Gauss' law for magnetism which may be written as

$$\vec{\nabla} \cdot \vec{B} = 0,$$

$$(53)$$

which is effectively a constraint on solutions of the magnetic field which must be satisfied for all time and negates the possibility for magnetic monopoles. In the continuous setting, given initial and boundary data which satisfy $\vec{\nabla} \cdot \vec{B} = 0$, the ideal MHD equations ensure a divergence-free evolution of the magnetic field. However, when a particular combination of spatial and temporal discretization is applied to the ideal MHD equations in more than one dimension, unphysical non-zero $\vec{\nabla} \cdot \vec{B}$ errors can emerge and pollute the solution. Numerical schemes designed to minimize these $\vec{\nabla} \cdot \vec{B}$ errors typically fall under

three categories: magnetic field projection [107]; divergence transport strategies [76,108]; and constrained transport schemes [109]. The first and second categories are used to remove small $\vec{\nabla} \cdot \vec{B}$ errors that arise from the non-zero discrete divergence of the discrete curl operator — a property that holds true in the continuous case. The third category, on the other hand, ensures $\vec{\nabla} \cdot \vec{B} = 0$ to machine accuracy for the specific discretization provided that the magnetic field is initially divergence-free for the same discretization procedure.

Standard one-dimensional solutions of the ideal MHD equations do not suffer from $\vec{\nabla} \cdot \vec{B}$ errors as the flux of the B_x component of the magnetic field is necessarily zero, resulting in a constant value for B_x which therefore satisfies the one-dimensional divergence-free property, provided that $B_x = $ constant is also satisfied by the initial data. However, in the context of data assimilation, $\vec{\nabla} \cdot \vec{B}$ errors can arise from the updates to the initial data inherent to the assimilation procedure. These errors go beyond those associated with the particular discretization scheme applied to the ideal MHD equations and occur regardless of the number of spatial dimensions for the problem. For instance, consider the EnKF analysis of Eq. (22) restricted to the three components of the magnetic field given by

$$\hat{\mathbf{B}}^{n,s} = \mathbf{B}^{n,s} + \mathbf{K}_{\mathbf{B}}^n(\mathbf{d}_{\mathbf{B}}^{n,s} - \mathbf{H}_{\mathbf{B}}^n\mathbf{B}^{n,s}), \tag{54}$$

as well as the preconditioned 4DVar update of the initial magnetic field given by

$$\mathbf{B}^0 = \mathbf{B}^b + \sqrt{\Sigma_{\mathbf{B}}^0}\mathbf{V}_{\mathbf{B}}, \tag{55}$$

where the $\mathbf{B}$ subscripts indicate a restriction to the magnetic field components only. The resulting assimilated magnetic fields will not be divergence free as both

$$\text{div}\,\mathbf{K}_{\mathbf{B}}^n(\mathbf{d}_{\mathbf{B}}^{n,s} - \mathbf{H}_{\mathbf{B}}^n\mathbf{B}^{n,s}) = \mathbf{0} \tag{56}$$

and

$$\text{div}\,\sqrt{\Sigma_{\mathbf{B}}^0}\mathbf{V}_{\mathbf{B}} = \mathbf{0} \tag{57}$$

do not generally hold. Here, "div" denotes the discrete divergence operator evaluated at each cell. As a consequence, data assimilation induced $\vec{\nabla} \cdot \vec{B}$ errors can become considerably large, resulting in unphysical magnetic fields and negative mass or pressure values (i.e., physically non-realizable solutions) which can prevent practical simulations in many cases. Effective $\vec{\nabla} \cdot \vec{B}$ treatments for data assimilation with global MHD models remains an open question and, for this reason, observational magnetic field data are therefore often disregarded to circumvent such problems in practical simulations. See for example, numerical DA experiments reported by Lang et al. [57].

To enable the assimilation of magnetic field data while systematically addressing the solenoidal constraint associated with the magnetic field, this study examines and quantitatively compares two strategies adapted from the numerical MHD literature. These techniques are further described in the subsections that follow, and numerical experiments are conducted in Section 8.5. Both magnetic field projection and divergence transport are employed to mitigate $\vec{\nabla} \cdot \vec{B}$ errors introduced during data assimilation. The former entails a moderately expensive additional correction step after each assimilation update, while the latter necessitates a modification to the ideal MHD numerical solution procedure. The constrained transport approach is excluded from this study due to its inability to correct initially unphysical magnetic fields.

In addition to the projection and divergence transport techniques, the large assimilation induced $\vec{\nabla} \cdot \vec{B}$ errors can be widely reduced in the variational approach by a judicious choice of the background covariance matrix. In particular, if the background covariance for the magnetic field is designed to have divergence-free sample paths, then Eq. (57) holds with some numerical noise [110], after which the projection or divergence transport techniques can more effectively minimize the remaining $\vec{\nabla} \cdot \vec{B}$ errors. The same approach is not effective with the

EnKF, in large part, due to covariance localization and the finite size of the ensemble.

7.1. Divergence transport of data assimilation induced errors

The main idea behind divergence transport strategies is to remove via transport any non-zero $\vec{\nabla} \cdot \vec{B}$ errors that arise from the discretization of the ideal MHD equations within the computational domain. A straightforward means to deal with data assimilation induced $\vec{\nabla} \cdot \vec{B}$ errors is to rely on the same divergence transport strategy. Although divergence transport was not originally intended for this application, this simple approach offers an effective baseline technique. Indeed, many ideal MHD codes already incorporate divergence transport, requiring no additional modifications.

Powell's 8-wave formulation [76] is considered here for divergence transport due to its prevalence in many ideal MHD codes as well as its relative simplicity. In Powell's method, the ideal MHD equations are written in symmetrizable form while retaining terms proportional to $\vec{\nabla} \cdot \vec{B}$, and are given by

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{F} = \mathbf{S}, \tag{58}$$

where $\mathbf{U}$ and $\mathbf{F}$ remain unchanged, and the 'source term' vector, $\mathbf{S}$, is given by

$$\mathbf{S} = -\nabla \cdot \vec{B}\begin{bmatrix} 0 \\ \vec{B} \\ \vec{u} \\ \vec{u} \cdot \vec{B} \end{bmatrix}. \tag{59}$$

Consequently, the semi-discrete form of the upwind finite-volume formulation applied to the ith computational cell is supplemented by a numerical source term and is given by

$$\frac{d\mathbf{U}_i}{dt} = -\frac{1}{\Delta x}(\mathcal{F}_{i+1/2} - \mathcal{F}_{i-1/2}) - \frac{B_{x,i+1} - B_{x,i-1}}{2\Delta x}\begin{bmatrix} 0 \\ \vec{B} \\ \vec{u} \\ \vec{u} \cdot \vec{B} \end{bmatrix}_i. \tag{60}$$

In order to account for the numerical source term, the discrete adjoint equations must also be appropriately modified. Direct differentiation followed by application of the transpose operator to the discretized Powell source vector, $\mathbf{S}$, yields the following expression given by

$$\sum_{j=-1}^{1}\left[\frac{\partial \mathbf{S}_{i+j}}{\partial \mathbf{U}_i}\right]^{\mathsf{T}}\lambda_{i+j}, \tag{61}$$

which must be added to Eq. (50) for $\hat{\lambda}_i$ given above. The corresponding expressions for each of the Jacobian matrices, $\partial\mathbf{S}_{i+j}/\partial\mathbf{U}_i$, are provided in Appendix C.

7.2. Divergence-free projection of magnetic field data

The so-called projection method, first introduced for MHD calculations by Brackbill and Barnes [107], is used to project numerically approximated magnetic fields with non-zero divergence to the space of divergence- free vector fields. The method is based on Helmholtz decomposition, which states that a vector field can be split into irrotational and solenoidal components under fairly weak conditions. Application of Helmholtz decomposition to a numerically approximated magnetic field, $\vec{B}^*$, yields

$$\vec{B}^* = \vec{\nabla} \times \vec{A} + \vec{\nabla}\phi, \tag{62}$$

where the first term is solenoidal, and the second is irrotational. Taking the divergence of $\vec{B}^*$ results in the following Poisson equation for ϕ,

$$\nabla^2\phi = \vec{\nabla} \cdot \vec{B}^*. \tag{63}$$

After solving, a purely solenoidal magnetic field, $\vec{B}$, can be obtained by subtracting the irrotational component, yielding

$$\vec{B} = \vec{B}^* - \vec{\nabla}\phi. \tag{64}$$

In the context of the data assimilation for ideal MHD considered herein, the projection method is applied to every ensemble member after each analysis step in the EnKF, or to each initial magnetic field iterate obtained during the 4DVar optimization procedure. The EnKF method therefore incurs a significantly larger computational cost penalty by employing the projection method compared to 4DVar approach. To maintain consistent gradient evaluations in the variational approach, a discrete adjoint of the Poisson solver associated with the magnetic field must also be implemented in the case of multidimensional computations. However, for the one-dimensional problems of interest herein, a simplified one-dimensional analog of the projection method can be formulated. It can be shown that the divergence-free projection of Eq. (64) is equivalent to solving the following minimization problem [111]

$$\underset{\vec{B}}{\text{minimize}} \qquad \|\vec{B} - \vec{B}^*\|^2, \tag{65a}$$

$$\text{subject to} \qquad \vec{\nabla} \cdot \vec{B} = 0. \tag{65b}$$

Replacing the constraint, $\vec{\nabla} \cdot \vec{B} = 0$, with its one-dimensional equivalent, $\partial B_x/\partial x = 0$, yields the following exact solution to the above minimization problem

$$B_x = \frac{1}{b-a} \int_a^b B_x^* dx, \tag{66}$$

where a and b are the spatial domain boundaries. It becomes immediately clear that the one-dimensional analog of the projection method is simply the average of the non-solenoidal component of the magnetic field. Similar to the multidimensional setting, the one-dimensional projection is also a global operation and to account for the projection in the 4DVar gradient evaluation, the B_x entries of λ^0 can be merely replaced by their corresponding average.

8. Data assimilation (DA) numerical experiments

Several numerical experiments are now described to demonstrate and compare the efficacy of the EnKF and 4DVar data assimilation algorithms presented in Sections 5 and 6 above. So-called twin experiments are considered for which synthetic observations were gathered from a reference simulation. To recover the reference simulation, the synthetic observations were assimilated with an erroneously initialized model. These types of numerical experiments are very standard in data assimilation studies as they allow for a controlled verification and comparison of the data assimilation algorithms of interest.

Three initial-value problems (IVPs) associated with the 1D ideal MHD equations are considered herein and, for each case, the sparsity of observations was varied. Twin experiment simulation results of a radially expanding plasma are first examined in Section 8.1. The canonical, MHD, shock-tube, problem of Brio and Wu [112] is then considered in Section 8.2, noting that this IVP is commonly used in the assessment, verification, and validation of shock-capturing, numerical, solution methods for ideal MHD. A third and final IVP, inspired by Shu and Osher's shock-tube problem [113,114] and mimicking shock-turbulence interaction, is also investigated in Section 8.3. For each IVP, the numerical solution was taken to be the "true" model evolution. Synthetic observations were then generated by sparsely sampling the conserved quantities and adding Gaussian noise. The background initial conditions were also iterated in time in order to quantify the effects of data assimilation in the reduction of errors compared to a simulation with no data assimilation. For the remainder of this paper, the term 'background' is used to refer both to the *a priori* initial conditions, as well as the unsteady solution quantities obtained by starting from these

initial conditions. Note however that the first set of assimilation results presented in Sections 8.1–8.3 do not include synthetic observations for the x-component of the magnetic field, B_x; this component of the solution did not participate in the assimilation process as any data-induced deviation in B_x from a constant value could potentially violate the solenoidal property of the magnetic field and introduce MHD solution errors without careful treatment. Additional numerical experiments of the proposed $\vec{\nabla} \cdot \vec{B}$ treatment strategies described in Section 7 are subsequently presented and discussed in Section 8.5 for which the sequential and variational data assimilation strategies are applied to all MHD variables, including B_x. Additional data assimilation results are provided in Section 8.4 which assess the sensitivity of the data assimilation procedures to the total number of model evaluations.

It should be noted that several simulations suffered from stability issues due to the occurrence of both negative values of the plasma mass and pressure brought about by the assimilated updates arising from the EnKF approach. These physically non-realizable solutions were largely mitigated by forcing all unphysical density and pressure values to be equal to minimum thresholds of 0.005. This ad-hoc technique was sufficient for performing the simulations in most situations; however, stability issues of this type were unavoidable in certain cases as described in the discussions to follow. Similar stability issues were not observed in the 4DVar results and the corresponding physical realizability corrections were not required. Not unexpectedly, the direct use of the numerical solution as represented by the discretized form of the ideal MHD equations within the constrained minimization problem for the initial data leads to more robust assimilated results that respect the physical realizability of the solutions (i.e., positivity of the plasma density and pressure) for the non-linear IVPs of interest here.

In an actual operational space weather forecasting setting, the background initial conditions, $\mathbf{U}^b$, would be provided by previous simulations and/or assimilation and would be consistent with the $\mathcal{N}(\mathbf{U}_{\text{truth}}^0, \boldsymbol{\Sigma}_{\mathbf{U}}^0)$ Gaussian distribution. Additionally, the specification of the model error covariance, $\boldsymbol{\Sigma}_{\mathbf{U}}^0$, would be an important endeavor in its own right. Nevertheless, in the idealized twin experiments of Cases I, II, and III considered herein, $\boldsymbol{\Sigma}_{\mathbf{U}}^0$ was specified directly and was formed in terms of a sample covariance of random curves as described by Evensen [25]. The background initial condition was then prescribed by a random sample drawn from $\mathcal{N}(\mathbf{U}_{\text{truth}}^0, \boldsymbol{\Sigma}_{\mathbf{U}}^0)$. This procedure permitted the direct comparison and evaluation of the EnKF and 4DVar data assimilation strategies in a consistent fashion. The observation error covariance matrix, $\boldsymbol{\Sigma}_{\mathbf{d}}^n$, was simply specified as a diagonal matrix containing the noise variance of the synthetic observations. In the EnKF experiments, an ensemble of 80 members was generated by sampling from $\mathcal{N}(\mathbf{U}^b, \boldsymbol{\Sigma}_{\mathbf{U}}^0)$. The mean of the ensemble corresponds exactly to the background initial conditions used in the 4DVar simulations. This allowed both data assimilation methods to be compared in an equal and fair manner. To keep the number of model evaluations equal in the two data assimilation strategies, the 4DVar scheme was limited here to 80 iterations of the optimization procedure. The resulting assimilated solutions were then assessed in terms of the root-mean-square (RMS) error at time step n, ϵ^n, given by

$$\epsilon^n = \sqrt{\sum_{i=1}^{M} \left\| (\mathbf{U}_i^n)_{\text{truth}} - \mathbf{U}_i^n \right\|^2}. \tag{67}$$

where M is the total number of cells within the 1D computational mesh. The latter is a measure of the discrepancy between the assimilated and true solutions of each IVP.

8.1. Case I : Plasma-expansion initial-value problem

An IVP associated with a plasma expansion process with fully differentiable initial conditions is first considered. This case, referred to as Case I, involves hyperbolic non-linear wave transport with different components of the solution propagating at rather widely different

velocities and consists of two symmetric rarefaction waves moving supersonically in opposite directions. The initial velocity field is given by

$$
u_x =
\begin{cases}
-\dfrac{2}{\gamma}, & x \leq -\dfrac{3}{2}, \\[6pt]
-\dfrac{1}{\gamma}\left(1 - \tanh\dfrac{x+1}{1/4-(x+1)^2}\right), & -\dfrac{3}{2} < x < -\dfrac{1}{2}, \\[6pt]
0, & -\dfrac{1}{2} \leq x \leq \dfrac{1}{2}, \\[6pt]
\dfrac{1}{\gamma}\left(1 + \tanh\dfrac{x-1}{1/4-(x-1)^2}\right), & \dfrac{1}{2} < x < \dfrac{3}{2}, \\[6pt]
\dfrac{2}{\gamma}, & x \geq \dfrac{3}{2},
\end{cases}
\tag{68}
$$

$$
u_y = u_z = 0,
\tag{69}
$$

with $\gamma = 2$. The sound speed, a, is specified according to

$$
u_x - \frac{2}{\gamma - 1}a = -\frac{2}{\gamma - 1}, \qquad x \leq 0,
\tag{70}
$$

$$
u_x + \frac{2}{\gamma - 1}a = \frac{2}{\gamma - 1}, \qquad x > 0.
\tag{71}
$$

The remaining hydrodynamic quantities are obtained using

$$
\rho = \gamma a^{\frac{2}{\gamma - 1}},
\tag{72}
$$

$$
p = \frac{\rho a^2}{\gamma},
\tag{73}
$$

which corresponding to the usual isentropic flow relations. Finally, the following initial magnetic field is imposed

$$
B_y =
\begin{cases}
\frac{1}{2}(\cos(2\pi(x+1)) + 1), & \frac{1}{2} < |x| < \frac{3}{2}, \\[6pt]
0, & \text{otherwise},
\end{cases}
\tag{74}
$$

$$
B_x = \frac{1}{2},
\tag{75}
$$

$$
B_z = 0.
\tag{76}
$$

The initial distributions of the density, x momentum, and y magnetic field component corresponding to the true and background solutions for the Case I plasma-expansion IVP are given in Fig. 2. Numerical simulations were performed for this case using a 1D domain, $-4\pi \leq x \leq 4\pi$, with a grid consisting of 1000 uniform computational cells. A localization length of 4 was used for the EnKF assimilations. The numerical time integration was performed in two phases: a data assimilation phase; and a forecast phase. The assimilation phase takes place from $t = 0$ to $t = 5$ with $N = 1250$ constant time steps. During this time window, the synthetic data was assimilated using all of the MHD solution variables except for the x-component of the magnetic field, B_x, which was assumed constant and equal to the exact value of $1/2$. As a first step, this avoided the creation of $\vec{\nabla} \cdot \vec{B} = 0$ constraint errors arising from the sequential and variational data assimilation procedures. The subsequent forecast phase takes place from $t = 5$ to $t = 10$ and also consists of 1250 constant time steps. In this phase, data is no longer ingested in the simulation. The assimilated solutions at the end of the previous phase are merely used as initial conditions during the forecast phase.

Synthetic observations are generated along the path of a fictitious observer with k "orbits" of the physical space domain, x. For example, the space–time locations of the observer corresponding to 5 orbits are depicted in Fig. 3(a). Observations corresponding to $k = 5, 10, 20$, and 40 were examined. Additionally, observations from two stationary observers were also generated and assimilated. Each stationary observer measured an eighth of the spatial domain near the ends of the channel. The space–time locations of the stationary observers are shown in Fig. 3(b). The conserved quantities were sampled from the true solution at the observer's space–time locations and Gaussian noise was added such that most of the noise (three standard deviations) was within 5% of the maximum value of the quantity. As the solution variables ρu_z and B_z have constant zero values, the noise statistics corresponding to ρu_x and B_y were used for ρu_z and B_z, respectively. The standard deviations

Table 2
Standard deviation of the components of solution noise used in synthetic observations for the Case I plasma-expansion initial-value problem.

	ρ	ρu_x	ρu_y	ρu_z	B_y	B_z	E
SD	0.033	0.014	0.005	0.005	0.017	0.017	0.019

of the added noise corresponding to each conserved solution quantity are summarized in Table 2.

The temporal evolution of the RMS error of the background and assimilated solutions is shown in Fig. 4. Additionally, a summary of the error reduction for each of the observation systems is summarized in Table 3. Both the EnKF and 4DVar assimilation techniques significantly reduced the error at the end of the assimilation window and the cumulative error in the forecast phase, with the 4DVar scheme outperforming the EnKF method. The assimilation involving observations near the domain boundaries indicates a substantial difference in the performance of the data assimilation algorithms. While the 4DVar DA method reduced the cumulative RMS error to 16.18% of the background error, the EnKF technique was only able to reduce the cumulative error to 53.24%. This large discrepancy is due to the hyperbolic nature of the governing ideal MHD equations and the resulting transport via wave motion associated with the ideal MHD description. With the EnKF approach, improvements brought about by the assimilation of observations only remain and propagate outwards with the rest of the solution content and rapidly leave the domain. Information is thus not able to be carried back into the middle of the domain for this IVP and the assimilated solution errors remain high. However, with the 4DVar scheme on the other hand, the adjoint equations allow information from observations in the outer portions of the 1D solution domain to be transmitted inwards towards the domain interior as dictated by the domain of dependence. As shown in the solution profiles of Fig. 5 and x-t density error plots of Fig. 6, while the EnKF solution recovers the true solution in the outer portions of the domain where the data assimilation is applied, the EnKF solution is very similar to the original background solution in the inner portions of the domain and the resulting RMS error is high. In contrast, the 4DVar assimilation procedure provides a clearly superior estimate that is very close to the true solution over the entire spatial domain with a far more uniform distribution of the RMS error.

The difficulties observed with the EnKF method considered here could potentially be alleviated through the use of both an ensemble Kalman smoother (EnKS) [81] and more sophisticated localization strategies. These techniques can help mitigate the rapid escape of state updates via wave transport phenomena. However, such methods were not considered here and the performance of these techniques is not expected to match that of 4DVar in the context of MHD outflows. In an ideal setting, observations would be localized both in space and time around characteristic lines intersecting with an analysis point [21]. For the complex and non-linear characteristics pertaining to the ideal MHD equations, the design of an EnKF or EnKS localization scheme that maintains these properties is a formidable challenge. In contrast, the adjoint equations of a 4DVar approach respect the inherent transport via wave propagation features of the underlying partial differential equations and naturally propagate information along the characteristic lines, allowing observations to correct the state in manner that respects their domain of dependence. This intrinsic feature of the 4DVar approach makes it particularly well-suited for handling the observation systems and flow physics relevant in many space plasmas applications.

8.2. Case II : Brio-Wu shock-tube initial-value problem

The second IVP, Case II, is the well-known Brio-Wu shock-tube problem [112] which consists of a high-density, high-pressure plasma and a low-density, low-pressure plasma of opposite polarity initially separated by a thin membrane. Once the membrane is removed at
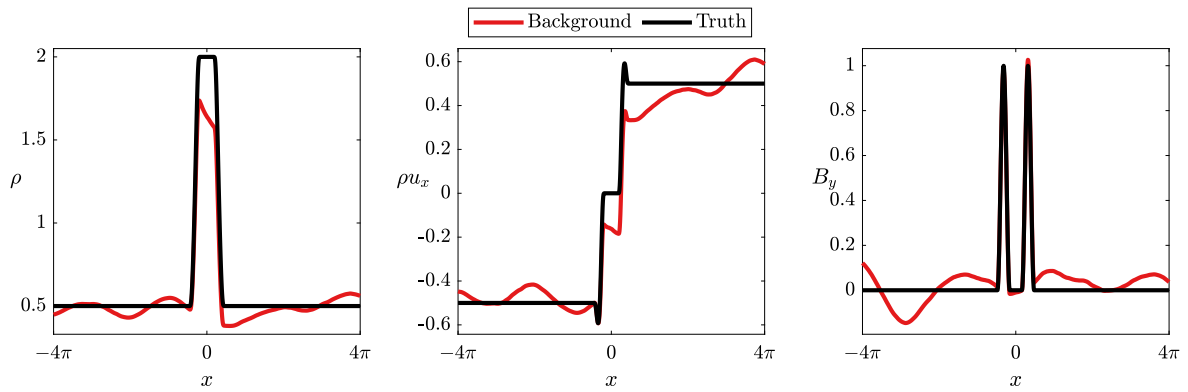
Fig. 2. Comparison of the mass density (left panel), x direction momentum (middle panel), and y direction magnetic field (right panel) corresponding to the true and background initial conditions for the Case I plasma-expansion initial-value problem.

(a)   (b)
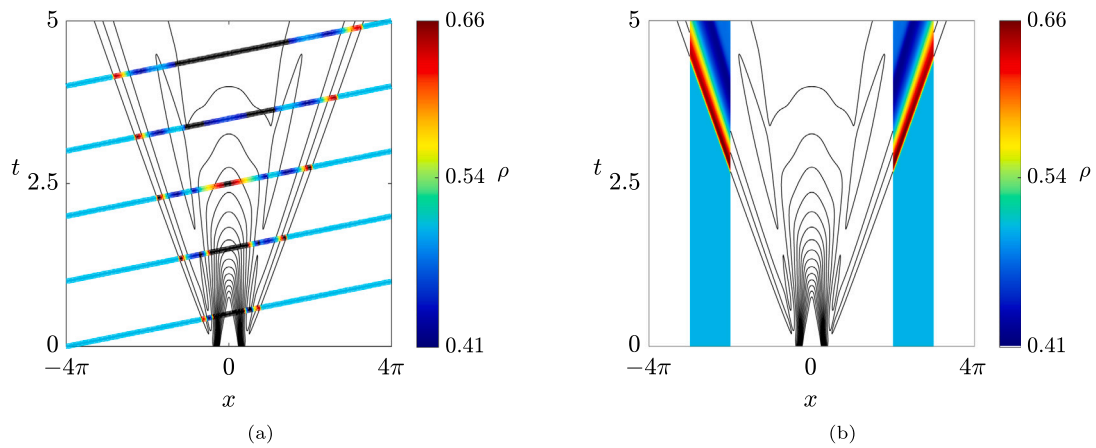
Fig. 3. Observation locations corresponding to (a) the orbiting observer with $k = 5$ orbits, and (b) the stationary observers, both for the Case I plasma-expansion initial-value problem. The observation locations are coloured by the true value of density. Observation-absent regions are denoted by the black contour plot of the space–time mas density profile.
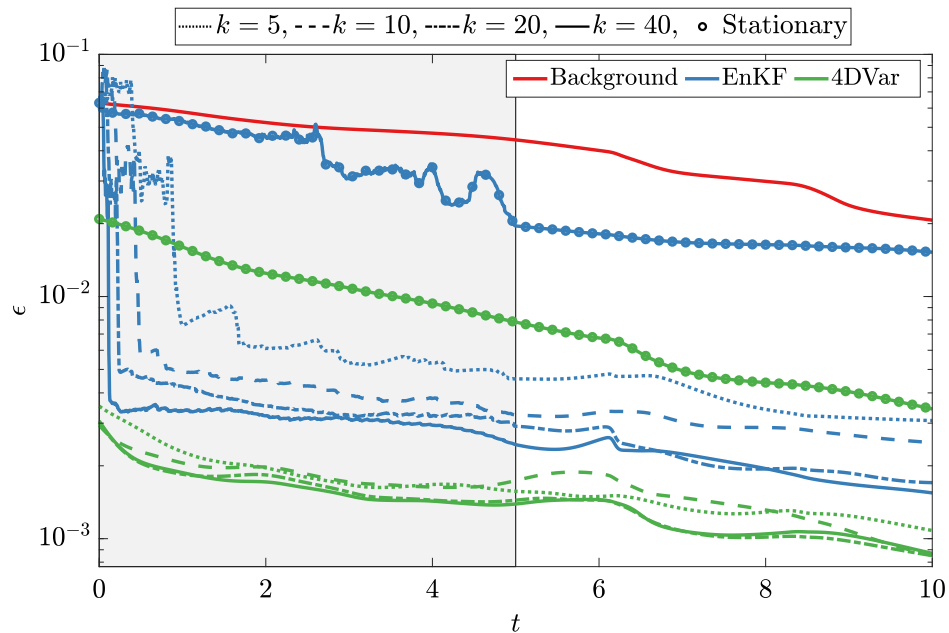
Fig. 4. Comparison of the RMS error, ϵ, in the background and assimilated estimates corresponding to the EnKF and 4DVar as a function of time, t, for the Case I plasma-expansion initial-value problem. The left half of the domain, shaded in grey, denotes the assimilation window. The forecast window corresponds to the right side. The RMS error is given for estimates corresponding to $k = 5, 10, 20,$ and 40 orbits of a synthetic observer, as well as stationary observers located near the boundaries of the 1D spatial domain.

Fig. 5. Predicted profiles of density (first row), x momentum (second row), y magnetic field (third row), and energy (fourth row) at the end of the assimilation phase (first column) and the forecast phase (second column) for the Case I plasma-expansion initial-value problem. The assimilated results correspond to the case with stationary observers.

Fig. 6. Predicted errors in the flow density for the background (left panel), EnKF (middle panel), and 4DVar (right panel) solutions as a function of position, x, and time, t, for the Case I plasma-expansion initial-value problem with observation from the stationary observers. Errors are computed according to $|\rho_t - \rho|/\rho_t$, where ρ_t is the true density. The assimilation and forecast time windows are separated by the horizontal line in the EnKF and 4DVar DA plots.
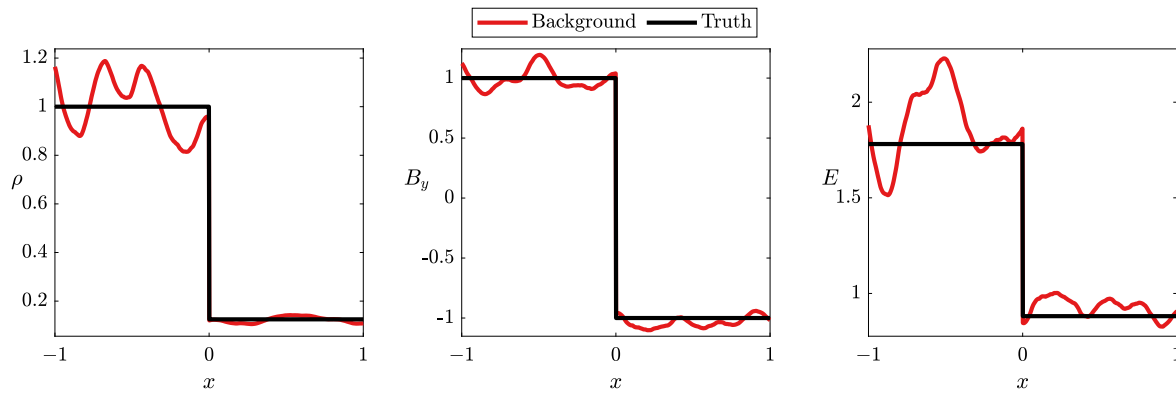
Fig. 7. Comparison of mass density (left panel), y direction magnetic field (middle panel), and energy (right panel) corresponding to the true and background initial conditions for the Case II Brio-Wu shock-tube initial-value problem.

Table 3
Relative computational error measured in terms of the RMS error at the end of the assimilation (2nd column) and sum of the RMS error in the forecast phase (3rd column) for the Case I plasma-expansion initial-value problem. Both error metrics are given as a percentage of the error in the background solution.

Observation system	$\epsilon^N / \epsilon^N_{\text{background}}$		$\sum_{n=N+1}^{2N} \epsilon^n / \sum_{n=N+1}^{2N} \epsilon^n_{\text{background}}$	
	EnKF	4DVar	EnKF	4DVar
$k = 5$	10.31%	3.54%	12.22%	4.21%
$k = 10$	7.31%	3.92%	9.30%	4.42%
$k = 20$	6.57%	3.25%	6.90%	3.56%
$k = 40$	5.51%	3.14%	6.45%	3.62%
Stationary	44.11%	17.69%	53.24%	16.18%

Table 4
Standard deviation of the components of solution noise used in synthetic observations for Case II Brio-Wu shock-tube initial-value problem.

	ρ	ρu_x	ρu_y	ρu_z	B_y	B_z	E
SD	0.017	0.007	0.018	0.007	0.017	0.017	0.032

Table 5
Relative computational error measured in terms of the RMS error at the end of the assimilation (2nd column) and sum of the RMS error in the forecast phase (3rd column) for Case II Brio-Wu shock-tube initial-value problem. Both error metrics are given as a percentage of the error in the background solution.

Observation system	$\epsilon^N / \epsilon^N_{\text{background}}$		$\sum_{n=N+1}^{2N} \epsilon^n / \sum_{n=N+1}^{2N} \epsilon^n_{\text{background}}$	
	EnKF	4DVar	EnKF	4DVar
$k = 5$	7.34%	3.23%	7.25%	3.85%
$k = 10$	5.04%	2.72%	5.31%	3.36%
$k = 20$	3.51%	2.53%	4.82%	2.99%
$k = 40$	3.00%	2.50%	2.88%	2.86%

$t = 0$, the solution consists of a number of strong non-linear propagating waves and discontinuities representing a range of MHD phenomena. A fast magnetosonic rarefaction wave, a slow compound wave, a contact discontinuity, a slow magnetosonic shock wave, and a fast magnetosonic rarefaction wave all result which propagate throughout the domain. The initial conditions for the Brio-Wu IVP is given by

$$\mathbf{W} = \begin{cases} [1, 0, 0, 0, \frac{3}{4}, 1, 0, 1]^{\mathsf{T}} & x \le 0 \\ [\frac{1}{8}, 0, 0, 0, \frac{3}{4}, -1, 0, \frac{1}{10}]^{\mathsf{T}} & x > 0 \end{cases}, \qquad (77)$$

with $\gamma = 2$. The true density, y magnetic field component, and energy are shown in Fig. 7 alongside the background initial condition. Both the true and background initial conditions were specified on a 1D domain,

$-1 \le x \le 1$, with a grid consisting of 800 uniform computational cells. A localization length of 0.15 was used for the EnKF assimilations. For the Case II IVP, both the assimilation and forecast phases of the simulation each consisted of $N = 2500$ constant time steps and 0.4 units of time.

As in Case I, synthetic observations were generated along the paths of a fictitious observer with $k = 5, 10, 20$, and 40 orbits. Synthetic noise was added to the observations according to the same methodology described for the Case I IVP. The standard deviations of the measurement noise are summarized in Table 4 for the Case II IVP. To avoid $\vec{\nabla} \cdot \vec{B} = 0$ constraint errors, the data assimilation was again applied using the synthetic data for all of the MHD variables except for the x-component of the magnetic field, B_x, which was assumed constant and equal to the exact value of $3/4$.

A comparison of the true, background, EnKF, and 4DVar predicted solutions at the end of the assimilation and forecast phases is shown in Fig. 8. The EnKF and 4DVar solution estimates are difficult to view as they lie extremely close to the truth or target solution. In order to better appreciate the errors in the assimilated solutions, the relative error in the density for the $k = 5$ Case II is given in Fig. 9 as a function of time, t. It is evident that both the EnKF and 4DVar methods significantly reduce the errors in the solution. The sequential nature of the EnKF can be seen by the sudden error reduction that takes place after the passage of the observer. In contrast, the error in the 4DVar solution estimate is initially low and remains so. It does not drastically change throughout the assimilation or forecast phase.

The results depicted in Fig. 9 correspond to the case with the fewest number of observations. As expected, further error reduction is achieved when a larger number of observations are assimilated as part of the simulation. The temporal evolution of the RMS solution error as a function of the spatial position with varying observer orbits, k, is provided in Fig. 10. Both the EnKF and 4DVar techniques significantly reduce the RMS error, with the 4DVar scheme showing overall better relative performance. Compared to the background simulation (i.e., foregoing data assimilation), the EnKF and 4DVar schemes reduced errors by up to 97% and 97.5% respectively. A detailed summary of the error reduction at the end of the assimilation window and cumulatively in the forecast phase is provided in Table 5. It can be seen that the errors remain low throughout the forecast phase, demonstrating the benefits of data assimilation for MHD-based forecasting. Notably, the 4DVar errors remain tightly bound despite the sparsity of observations, whereas the EnKF solution errors exhibit greater variance across the different observer configurations. Note that the errors in both the EnKF and 4DVar estimates exhibit sudden reductions in error at approximately $t = 0.7$. By comparing Figs. 9 and 10, it is evident that much of the predicted solution error is associated with the right-propagating fast rarefaction wave. Once this wave structure leaves the domain at $t \approx 0.7$, this source of error is no longer present and the overall RMS error is suddenly reduced.
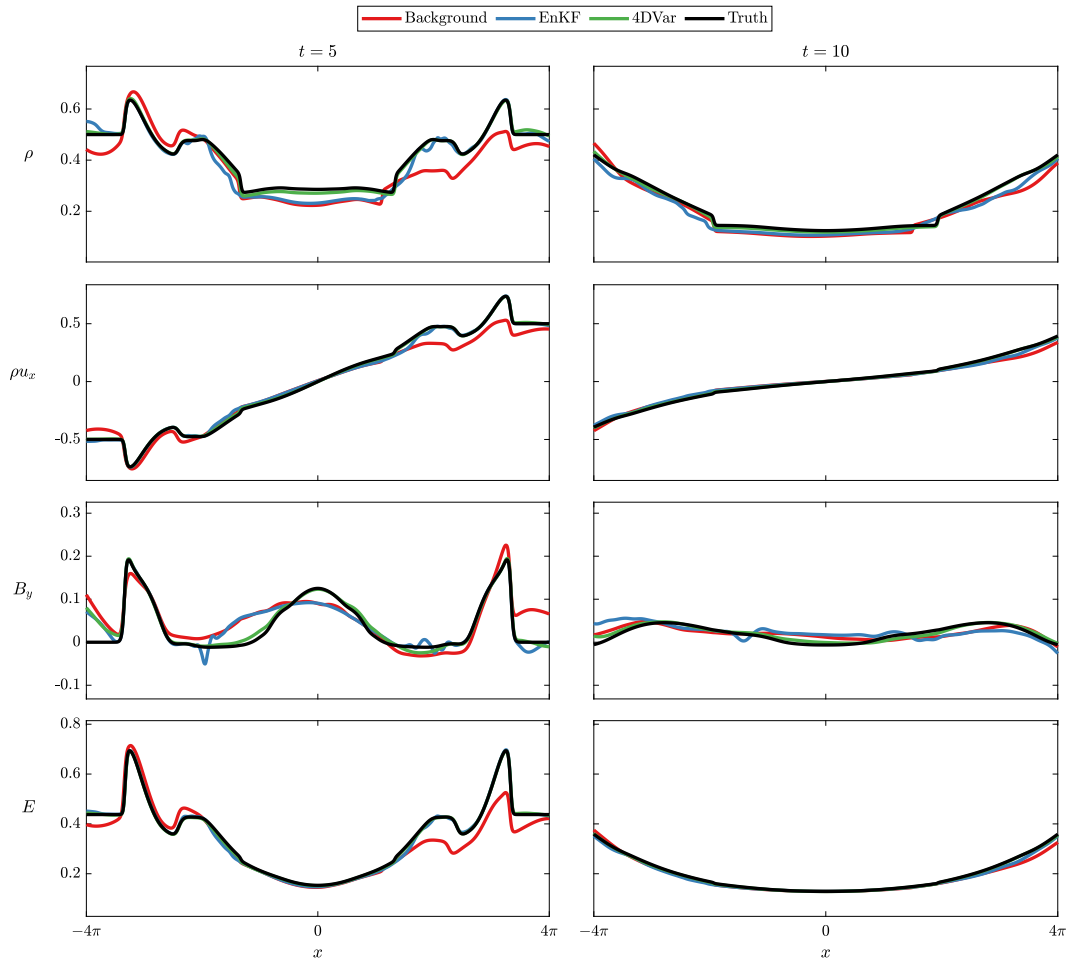
Fig. 8. Predicted profiles of density (first row), x momentum (second row), y magnetic field (third row), and energy (fourth row) at the end of the assimilation phase (first column) and the forecast phase (second column) for Case II Brio-Wu shock-tube initial-value problem. The assimilated results correspond to the observation system with 5 orbits.
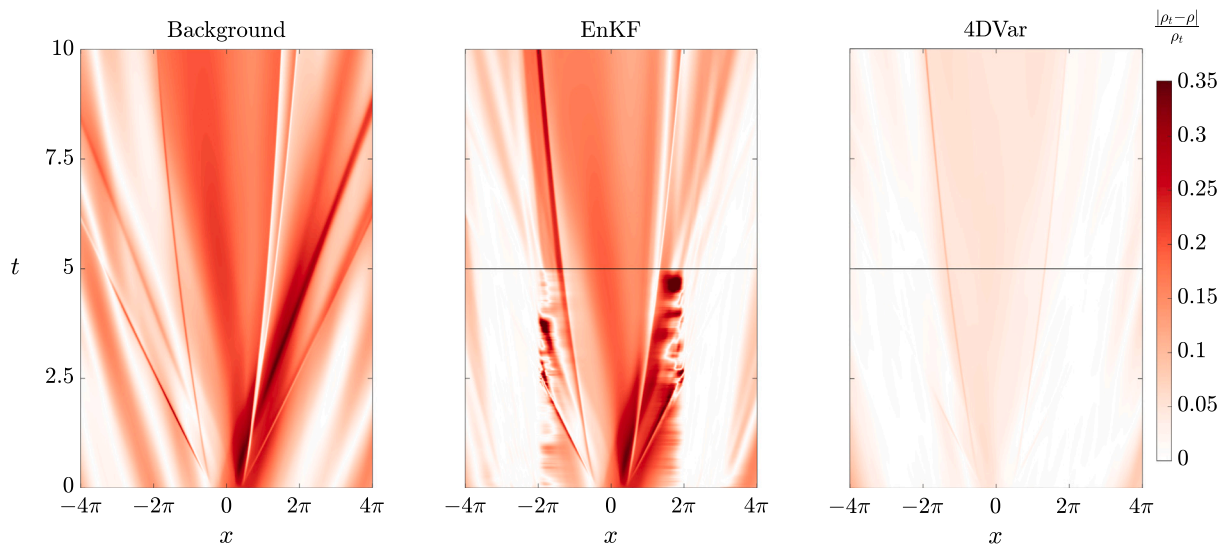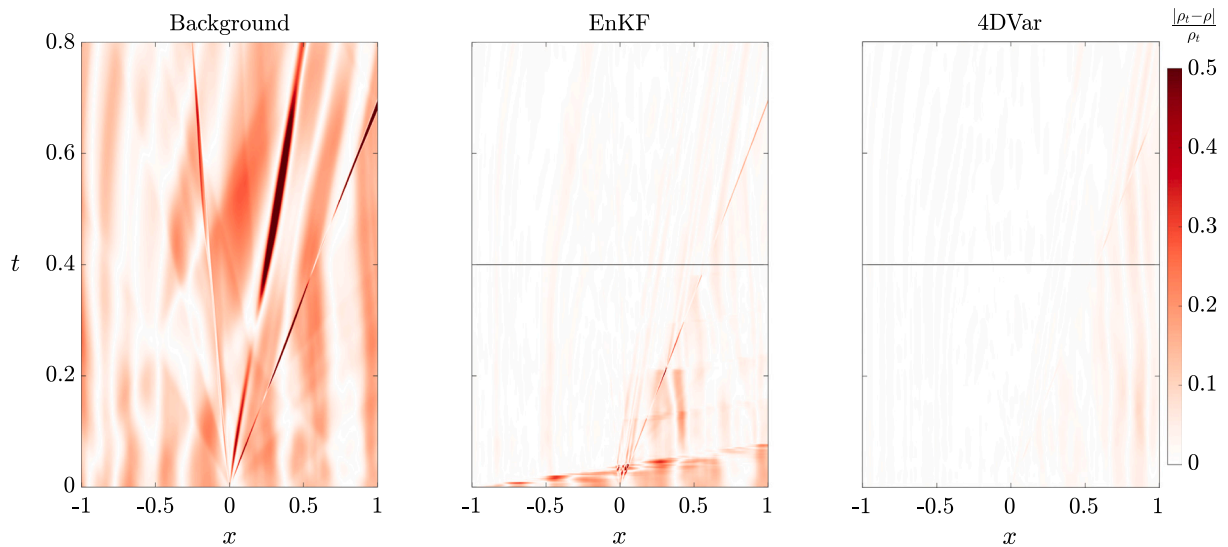

Fig. 9. Predicted errors in the flow density for the background (left panel), EnKF (middle panel), and 4DVar (right panel) solutions as a function of position, x, and time, t, for the Case II Brio-Wu, shock-tube initial-value problem with the $k = 5$ observation system. Errors are computed according to $|\rho_t - \rho|/\rho_t$, where ρ_t is the true density. The assimilation and forecast time windows are separated by the horizontal line in the EnKF and 4DVar DA plots.

Fig. 10. Comparison of the RMS error, ϵ, in the background and assimilated estimates corresponding to the EnKF and 4DVar as a function of time, t, for the Case II Brio-Wu shock-tube initial-value problem. The left half of the domain, shaded in grey, denotes the assimilation window. The forecast window corresponds to the right side. The RMS error is given for estimates corresponding to $k = 5, 10, 20$, and 40 orbits of a synthetic observer.

8.3. Case III: Shu-Osher inspired shock-turbulence interaction initial-value problem

The third and final IVP, Case III, consists of a shock-wave impinging upon a plasma with density fluctuations resulting in the formation of both magnetoacoustic and entropy waves downstream of the shock. This problem can be thought of as an idealization of highly non-linear shock-turbulence interactions, which can arise, for instance, in interplanetary shocks. A similar IVP was first suggested by Shu and Osher [113] to test high-order numerical schemes for gas dynamics and was later adapted by Susanto [114] for ideal MHD. Here, the IVP is mildly modified so that the density fluctuations drift towards the shock wave rather than remaining stationary. The initial conditions for Case III are thus given by

$$\mathbf{W} = \begin{cases} [3.5, 0, 0, 0, 1, 3.6359, 0, 42.0267]^\mathsf{T} & x \leq 1 \\ [\rho_R, -5.8846, -1.1198, 0, 1, 1, 0, 1]^\mathsf{T} & x > 1 \end{cases}, \tag{78}$$

where,

$$\rho_R = 1 + 0.2 \sin(5x), \tag{79}$$

and $\gamma = \frac{5}{3}$. The initial conditions of the true and background density, y momentum, and y magnetic field component corresponding to the Case III IVP are given in Fig. 11. Numerical calculations were performed on a 1D domain, $-4.5 \leq x \leq 4.5$, with a grid consisting of 800 uniform computational cells. Due to the highly non-linear and discontinuous nature of the plasma flow, a relatively short localization length of 0.1 was used for the EnKF assimilations. Moreover, stability issues were encountered with higher correlation lengths. The assimilation and forecast phases each consisted of $N = 750$ constant time steps and 0.6 units of time, respectively.

In contrast with the numerical experiments of Case I and II, the strong non-linear features of the modified Shu and Osher IVP for ideal MHD presented a significant challenge for the EnKF approach. In particular, the possible locations of synthetic observations were significantly restricted due to unphysical negative density and pressure values arising from the EnKF assimilation of these observations near the shock wave. These problems were caused by spurious correlations, as well as large magnitude differences, between pre- and post-shock quantities

Table 6
Standard deviation of the components of solution noise used in synthetic observations for Case III shock-turbulence interaction initial-value problem.

	ρ	ρu_x	ρu_y	ρu_z	B_y	B_z	E
SD	0.073	0.12	0.1	0.017	0.066	0.017	1.36

even when employing very small correlation lengths. As a result, the current EnKF method was rendered untenable for the assimilation of observations in close proximity to the shock-wave. In order to fairly compare the sequential and variational methods, synthetic observations were therefore located along the path of a fictitious observer that lagged the shock front by 1 unit in space, and measured a width of 0.1 units, as depicted in Fig. 12. Additionally, the standard deviations of the synthetic measurement noise are summarized in Table 6 for the Case III IVP. As for Cases I and II, the data assimilation was applied here using the synthetic data for all of the MHD variables except for the x-component of the magnetic field, B_x, which was assumed constant and equal to the exact value of 1, thus avoiding the creation of $\vec{\nabla} \cdot \vec{B}$ constraint errors.

Selected data assimilated solutions obtained using the EnKF and 4DVar methods are depicted in Fig. 13 alongside the true and background states. In addition, the temporal evolution of the RMS error for the assimilated and background simulations are provided in Fig. 14, and summarized in Table 7. The RMS error at the end of the assimilation phase was reduced by 4.13% and 38.86% compared to the background state RMS error with the EnKF and 4DVar approaches, respectively. However, in the early stages of the forecast phase–when observations are no longer available for assimilation–the RMS error in the EnKF solution surpasses the RMS error in the background state, while RMS errors in the 4DVar analysis remain small. Correspondingly, upon completion of the forecast phase, the RMS error achieved with the EnKF method *increased* by 88.56% compared to the background, while a similar error reduction of 35.95% was maintained in the 4DVar forecast.

The deteriorated performance of the EnKF method can be appreciated by inspection of Fig. 13. At the end of the assimilation phase ($t = 0.6$), plasma quantities generally match the true solution, with the
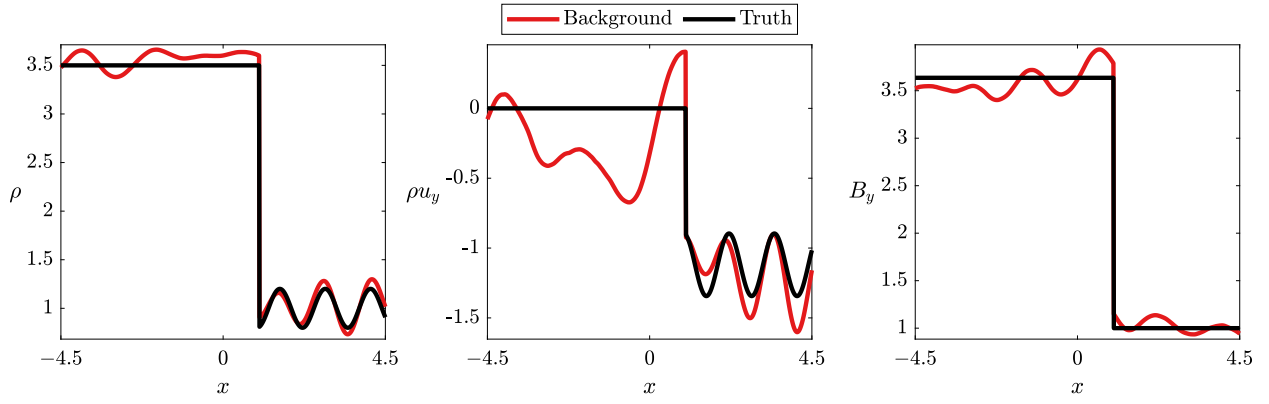
Fig. 11. Comparison of mass density (left panel), y direction momentum (middle panel), and y direction magnetic field (right panel) corresponding to the true and background initial conditions for the Case III shock-turbulence interaction initial-value problem.
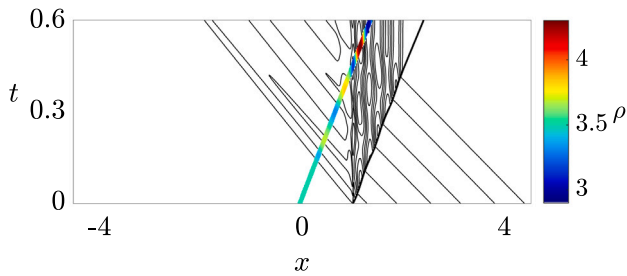
Fig. 12. Space–time locations of synthetic observations assimilated in the numerical experiments of the Case III shock-turbulence interaction initial-value problem. The observation locations are coloured by the true value of density. Observation-absent regions are denoted by the black contour plot of the space–time density profile.

Table 7
Relative computational error measured in terms of the RMS error at the end of the assimilation (2nd row) and sum of the RMS error in the forecast phase (3rd row) for Case III shock-turbulence interaction initial-value problem. Both error metrics are given as a percentage of the error in the background solution.

Error metric	EnKF	4DVar
$\epsilon^N / \epsilon^N_{\text{background}}$	95.87%	61.14%
$\sum_{n=N+1}^{2N} \epsilon^n / \sum_{n=N+1}^{2N} \epsilon^n_{\text{background}}$	188.56%	64.05%

exception of large errors introduced in the region spanned by entropy waves (the high-frequency fluctuations downstream of the shock). By contrast, significant errors are present throughout the domain at the end of the forecast phase ($t = 1.2$) due, in part, to a considerable underprediction of wave speeds. For instance, both the shock-wave and induced magnetoacoustic waves are significantly delayed compared to the true solution. These errors are brought about by the limited coverage of synthetic data, the short correlation length needed to maintain positive density and pressure, as well as the highly non-linear nature of the ideal MHD equations.

8.4. Sensitivity to the number of model evaluations

In the preceding DA simulation results, the ensemble size and maximum number of iterations were set to 80 as an upper limit to the number of model evaluations that can practically be considered in large scale calculations. The effect of a smaller limit on the number of model evaluations is now also examined here. The RMS error in the estimates of the solution at the end of the assimilation window as a percentage of the background RMS error for the three IVPs examined above is shown in Fig. 15 as a function of the number of MHD model evaluations. The RMS errors reported in Fig. 15 correspond to observation systems with 40 orbits of the observer for both Case I and Case II. For Case III, RMS errors corresponding to the shock following observation system are reported instead. The number of model evaluations was incremented by 10, starting with 20 and ending with 80. However, the minimum number of ensemble members in the EnKF was limited by the occurrence of highly unphysical solution content. Successful simulations could not be obtained with an ensemble size below 40, 30, and 75 in Case I, Case II, and Case III respectively. Additionally, if less data is ingested (as in the $k = 5$, 10, or 20 observation systems) a higher ensemble size was generally needed. When the ensemble was too small the statistics of the model were poorly estimated which in turn lead to unstable assimilation. On the other hand, while reducing the number of 4DVar iterations results in less optimal assimilation results, the estimated solutions always remained stable and physical.

As can be seen in Fig. 15, the RMS error for the 4DVar and EnKF DA methods are comparable when 60 or more model evaluations are used for the Case II, Brio-Wu IVP. In cases where fewer than 60 model evaluations are used, the 4DVar assimilations significantly outperformed the EnKF. For the Case I plasma-expansion IVP, and Case III shock-turbulence interaction IVP, the 4DVar technique achieved substantially greater error reduction than the EnKF approach for all values of the number of model evaluations considered. Unexpectedly however, in Case III, the RMS error increased with the number of model evaluations after 50 iterates, which is suggestive of an excessively long assimilation window for the 4DVar approach. As the number of model evaluations can be used here as a proxy for the overall computational cost of the DA procedure and the results of Fig. 15 suggest that, for the same computational cost, higher error reduction was achieved using the 4DVar approach compared to the EnKF method. However, it should be noted that parallel implementation of the EnKF DA method (not considered here) is rather trivially achieved, whereas parallel implementation of the 4DVar approach is generally more involved. Parallel implementation of the assimilation procedures can potentially reduce the time to solution considerably.

8.5. Data assimilation with all magnetic field components and $\vec{\nabla} \cdot \vec{B} = 0$ enforcement

The preceding numerical experiments presented in Sections 8.1, 8.2, and 8.3 do not involve the direct assimilation of synthetic data for the x-component of the magnetic field, B_x. This was to avoid possible non-zero $\vec{\nabla} \cdot \vec{B}$ errors introduced by the data assimilation corrections. Additional twin experiments are now considered for the same three IVPs outlined in Sections 8.1, 8.2, and 8.3 above; however, in these
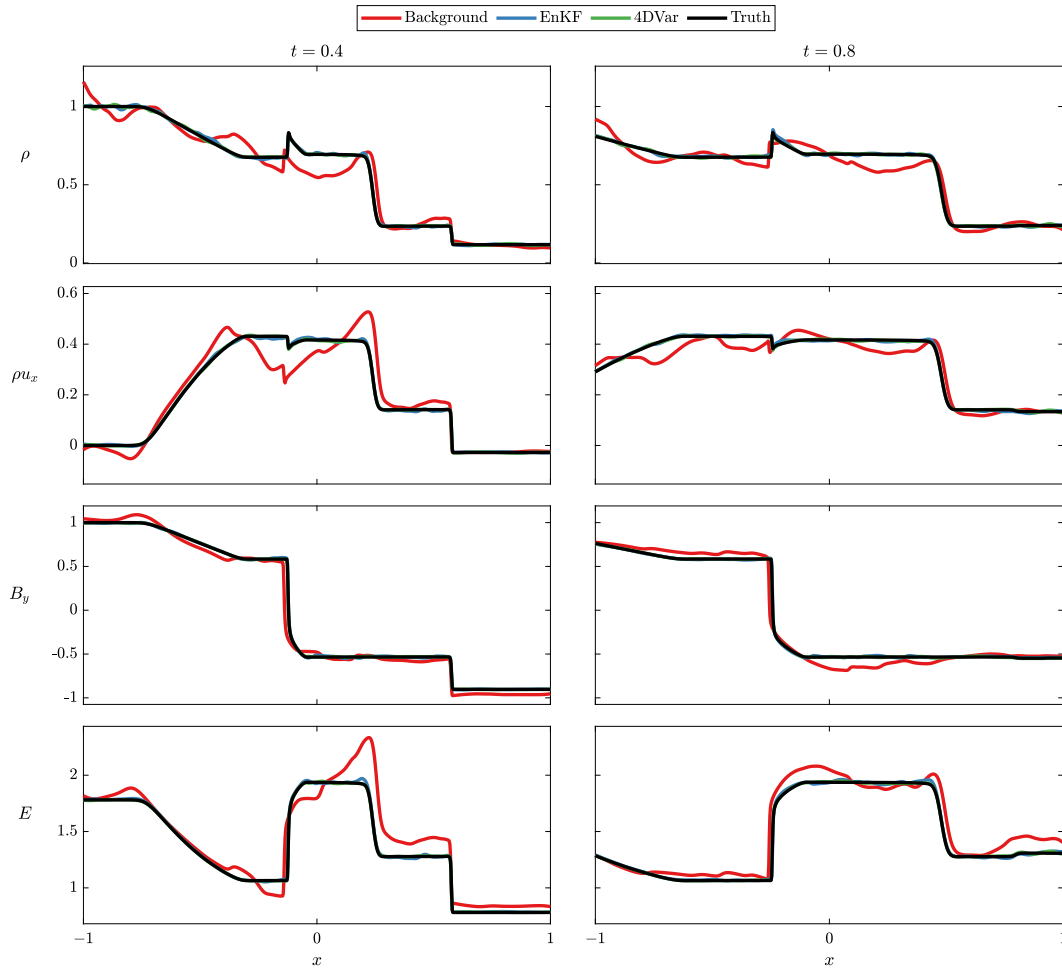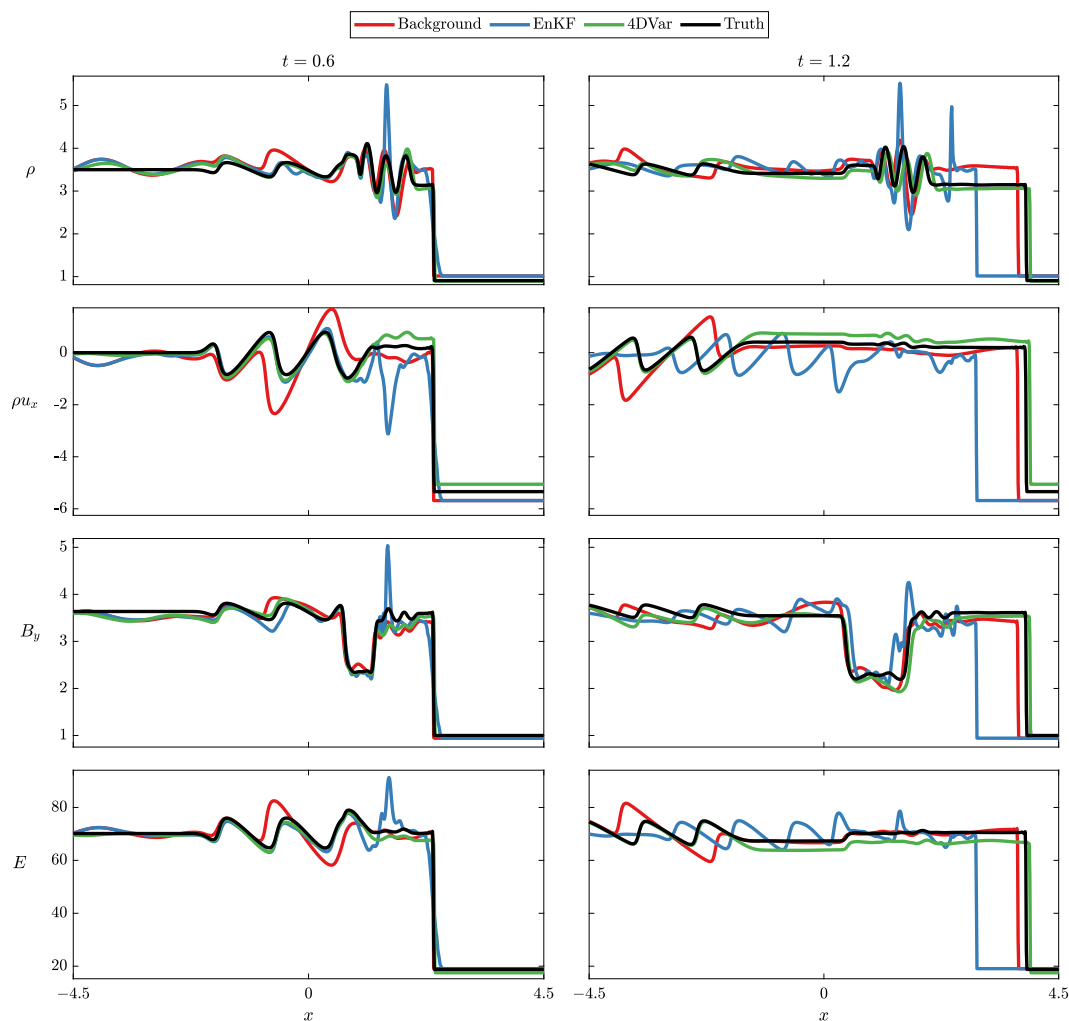
Fig. 13. Predicted profiles of density (first row), x momentum (second row), y magnetic field (third row), and energy (fourth row) at the end of the assimilation phase (first column) and the forecast phase (second column) for Case III shock-turbulence interaction initial-value problem.
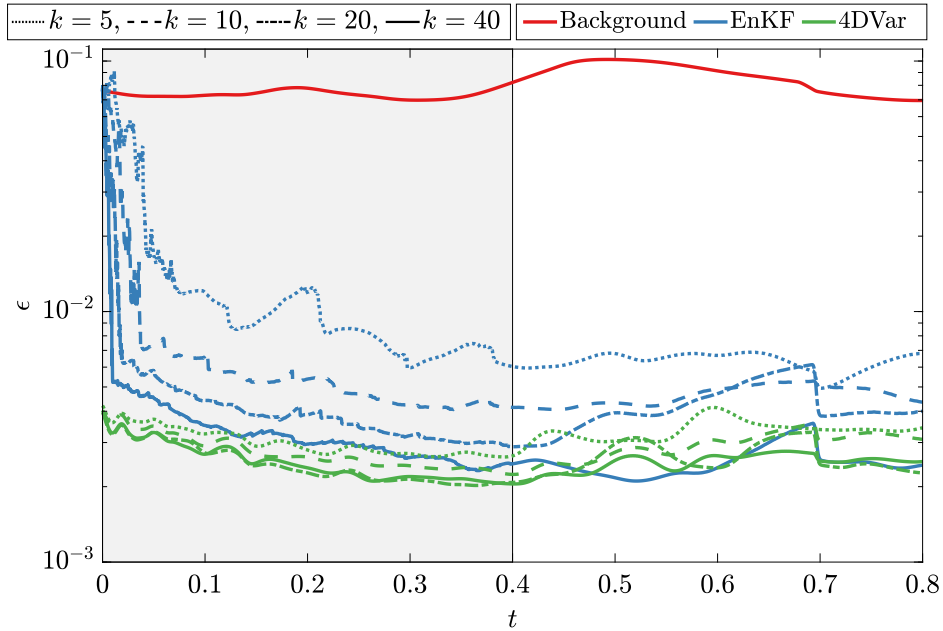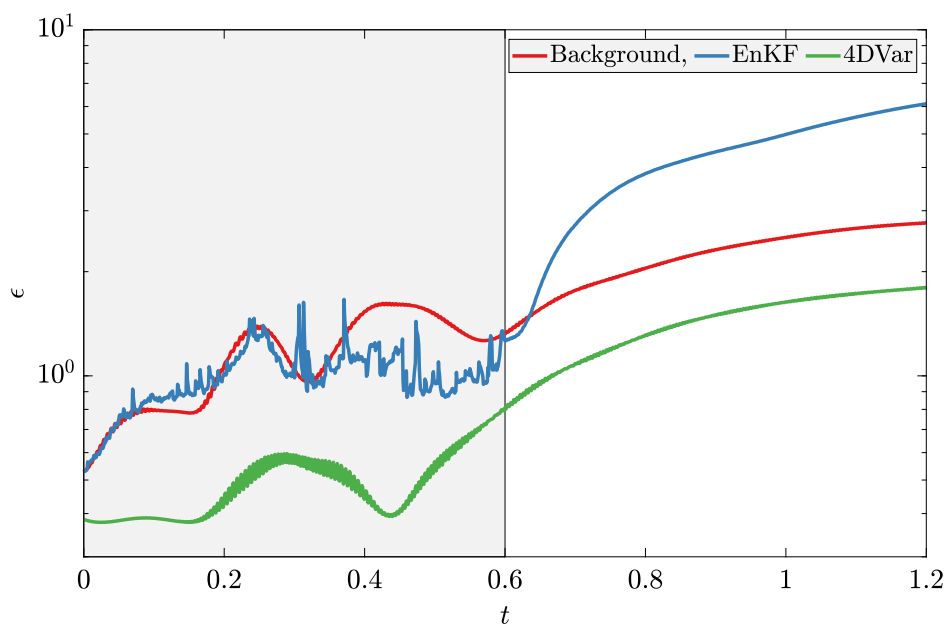

Fig. 14. Comparison of the RMS error, ϵ, in the background and assimilated estimates corresponding to the EnKF and 4DVar as a function of time, t, for the Case III shock-turbulence interaction initial-value problem. The left half of the domain, shaded in grey, denotes the assimilation window. The forecast window corresponds to the right side.
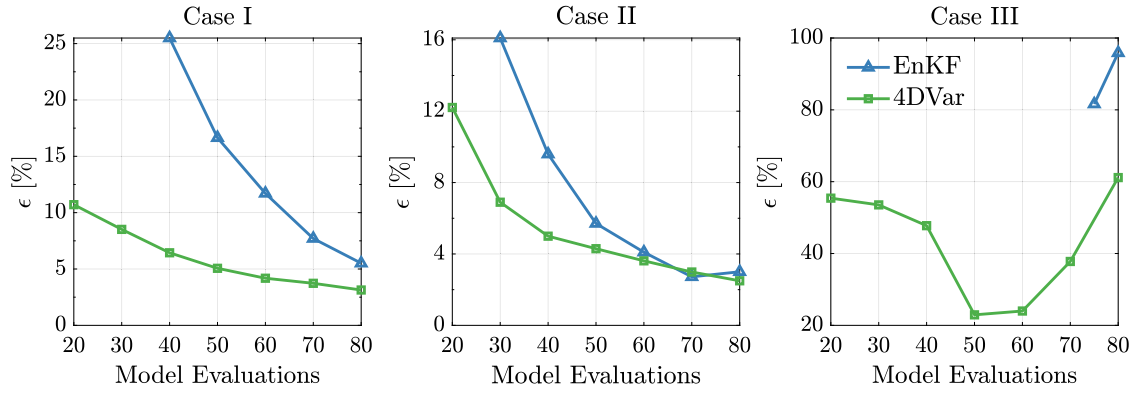
Fig. 15. Comparison of the 4DVar and EnKF RMS error at the end of the assimilation window as a percentage of the background RMS error. Data corresponds to the $k = 40$ observation system for both and Case I plasma-expansion (left panel) and Case II Brio-Wu shock-tube (middle panel) initial-value problems, as well as the shock following observation system of the Case III shock-turbulence interaction (right panel) initial-value problem.

Table 8
Relative computational error measured in terms of the RMS error at the end of the assimilation (2nd row) and sum of the RMS error in the forecast phase (3rd row) for the initial-value problems of Case I plasma-expansion; Case II Brio-Wu shock-tube; and Case III shock-turbulence interaction, each with all components of the magnetic field participating in data assimilation. Error metrics are given for the two $\vec{\nabla} \cdot \vec{B}$ error mitigation strategies considered in this paper–Powell's 8-wave ideal MHD formulation, and divergence-free projection of the magnetic field.

Error metric	$\vec{\nabla} \cdot \vec{B}$ Strategy	Case I		Case II		Case III	
		EnKF	4DVar	EnKF	4DVar	EnKF	4DVar
$\epsilon^N / \epsilon^N_{\text{background}}$	Powell	56.66%	5.89%	3.00%	2.25%	81.69%	82.64%
	Projection	27.57%	5.89%	2.45%	2.25%	–	92.83%
$\sum_{n=N+1}^{2N} \epsilon^n / \sum_{n=N+1}^{2N} \epsilon^n_{\text{background}}$	Powell	63.59%	3.90%	3.69%	3.06%	190.30%	69.61%
	Projection	23.28%	3.90%	2.65%	3.06%	–	80.40%

additional DA simulations, all components of the magnetic field are directly assimilated using the two strategies outlined in Section 7 so as to mitigate data assimilation induced $\vec{\nabla} \cdot \vec{B}$ errors. For these additional cases, the background initial conditions remain essentially unchanged except that the background solution for B_x is now also drawn from a random perturbation and is therefore not equal to the true exact solution for B_x in each case. Similarly, synthetic ensembles are now also generated to include variations in B_x. It should also be noted that a different observation system was considered for each of the three IVPs. For the plasma-expansion IVP of Case I, an observation system consisting of two stationary observers was used as described in Section 8.1. An observation system with $k = 20$ orbits was selected for the Case II Brio-Wu IVP. Finally, for the shock-turbulence interaction IVP of Case III, the shock-following observation system described in Section 8.3 was again adopted.

The temporal evolution of the RMS error in the EnKF and 4DVar estimates is shown in Fig. 16, alongside the background state, with each of the two $\vec{\nabla} \cdot \vec{B}$ error mitigation strategies considered here: the Powell 8-wave ideal MHD formulation and divergence-free projection of the magnetic field. In addition, crucial error metrics from each of the twin experiments are summarized in Table 8. The EnKF estimates for the first two IVPs showed higher performance when employing the projection approach for $\vec{\nabla} \cdot \vec{B}$ error mitigation in comparison to the Powell method. This performance difference should be expected as the global nature of the projection method yields a solenoidal magnetic field within the entire computational domain, even with sparse observations. This behaviour is also highlighted in the results for the x-component of the magnetic field obtained for the Case I twin experiments and depicted in Fig. 17. Use of the projection method results in an ideal MHD solution with a constant and steady value for B_x that closely matches the ground truth or exact solution. In contrast, the use of Powell's approach, further exacerbated by ensemble localization, produces an unsteady

and non-constant B_x profile, where the data-assimilated corrections are effectively transported away and leave the domain. It should be noted, however, that the use of the projection method in conjunction with the EnKF requires a projection operation at each observation time, which, in the case of large-scale multidimensional simulations, can be computationally demanding. Unfortunately, the aforementioned benefits of the projection method could not be realized in the case of the third IVP due to negative pressure issues with the EnKF approach.

With regard to the 4DVar estimates, very similar performance was achieved with both $\vec{\nabla} \cdot \vec{B}$ error mitigation strategies. For the first two IVPs, virtually identical assimilations were obtained with the Powell and projection methods. In the third IVP, Powell's method achieved marginally better performance. The similarity in performance between the Powell and projection approach can be attributed to the specification of an appropriate background error covariance for the magnetic field as described in Section 7 above. By constructing the background covariance matrix so that magnetic field perturbations are divergence free, the Powell and projection approaches need only mitigate the presence of small $\vec{\nabla} \cdot \vec{B}$ errors resulting from numerical noise. The effects of this approach are not only numerically advantageous, but also physically desirable. Comparing the EnKF and 4DVar estimates, the 4DVar predictions sustained significantly higher forecast accuracy compared to those of the EnKF method for the Case I and III IVPs. For Case II with the sparse data or least observations, the EnKF approach showed both higher and lower forecast accuracy compared to 4DVar approach when combined with the projection method and Powell approach, respectively.

9. Conclusions and future research

In the present study, both sequential and variational data assimilation strategies have been applied to constrain the evolution of
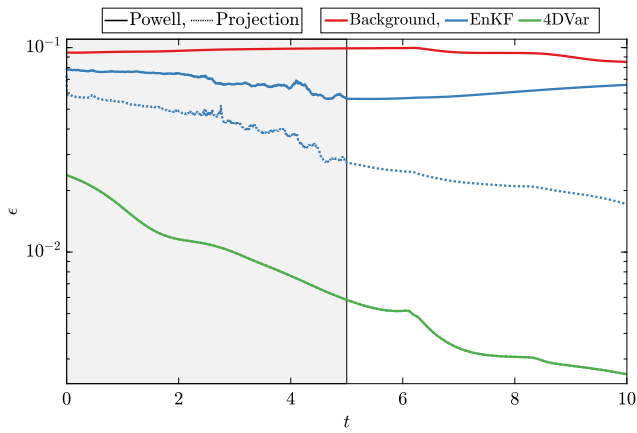
(a)
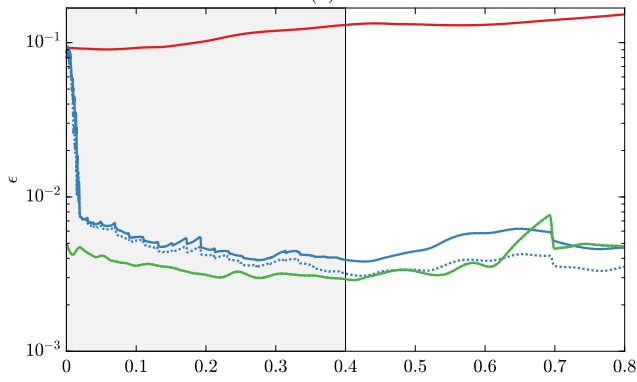
(b)
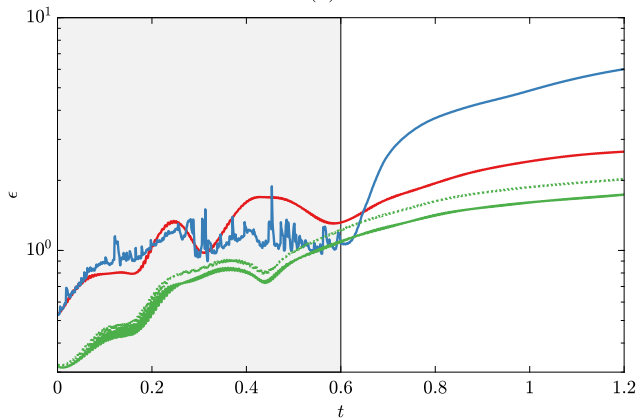
(c)

Fig. 16. Comparison of the RMS error, ϵ, in the background and assimilated estimates corresponding to the EnKF and 4DVar as a function of time, t, for the initial-value problems of (a) Case I plasma-expansion; (b) Case II Brio-Wu shock-tube; and (c) Case III shock-turbulence interaction, each with all components of the magnetic field participating in data assimilation. The left half of the domain, shaded in grey, denotes the assimilation window. The forecast window corresponds to the right side. RMS error is given for the two $\vec{\nabla} \cdot \vec{B}$ error mitigation strategies considered in this paper–Powell's 8-wave ideal MHD formulation, and divergence-free projection of the magnetic field. The 4DVar RMS error curves in (a) and (b) are visually indistinguishable due to close numerical proximity. In (c), the EnKF RMS data with the projection method is not present as a solution could not be obtained.

Fig. 17. Background and EnKF estimates of B_x in the left half of the physical domain for the Case I plasma-expansion initial-value problem with all components of the magnetic field participating in data assimilation. B_x profiles are given for the two $\vec{\nabla} \cdot \vec{B}$ error mitigation strategies considered in this paper–Powell's 8-wave ideal MHD formulation, and divergence-free projection of the magnetic field. The B_x estimates obtained with the Powell $\vec{\nabla} \cdot \vec{B}$ strategy is shown for several time steps throughout the forecast phase.

corrected by the assimilation of noisy and sparse synthetic observations of the plasma flow properties. Three highly non-linear IVPs were examined, the first involving smooth and continuous initial data, the second involving discontinuous initial conditions, and the third involving the propagation of a strong shock-wave in the presence of entropy fluctuations, with all three being quite representative of the class of unsteady MHD flows arising in heliospheric, solar wind, and geospace environment. For the three initial value problems, the locations of the synthetic observations were varied and their influence on the assimilated predictions were investigated. Careful head-to-head comparisons and assessments of the sequential and variational method simulation results were performed. In all situations, the 4DVar assimilation results were observed to surpass the error reduction that was obtainable with the corresponding EnKF approach. The most significant difference in performance occurred with the observation systems containing sparse data. For the observation system in which the synthetic observations were only made available near the boundaries of the spatial 1D domain, the EnKF solution corrections associated with the observational information was more local in nature and could only be transported towards the boundaries and out of the domain according to the transport of solution content by wave motion contained in the hyperbolic governing equations of the ideal MHD description. In contrast, via the inclusion of the solution of the corresponding adjoint equations for the dual problem, solution corrections arising from the observational information were also propagated inwards by the 4DVar assimilation procedure, allowing the plasma flow field to be corrected globally to a much higher degree. This non-local attribute of the variational approach is deemed to be a rather significant advantage in data assimilation applied to space weather forecasting, for which the observations are expected to be both sparse and remote from each other and the areas of interest. Furthermore, due to the super-magnetosonic and super-Alfvenic nature of the solar wind, variational data assimilation would allow solution corrections based on downstream observational data within the outer heliosphere to be carried back to the solar surface via the adjoint equations thereby improving the estimates of the plasma flow as also pointed out by Lang and Owens [33].

Difficulties associated with the positivity of pressure and density fields were also encountered with the EnKF based assimilation of observational data in proximity to strong shock-wave discontinuities, and

fully-ionized plasma flows associated with 1D initial value problems governed by the ideal MHD equations. The latter is a mathematical description that is commonly used in global MHD models for space weather forecasting. Poorly known initial conditions were successfully

thereby artificially limiting the location of observational data that could be ingested. This problem could potentially be addressed by localizing in a directional fashion, removing all correlations in the solution across shock discontinuities. This idea, however, was not pursued herein. In contrast, the same issues were not encountered with the variational approach, even for the strongly-nonlinear IVPs examined herein. By directly constraining the assimilated solutions to satisfy the underlying discretized solutions of the ideal MHD equations, physical realizability results are more readily achieved. This would appear to be an important and key benefit of the 4DVar approach for non-linear problems and provides further encouragement for the use of a variational approach in future, more complex, plasma, simulations.

The effects of the ensemble size for the EnKF approach and the maximum number of optimization iterations allowed in the 4DVar minimization procedure were also studied and assessed here. When the number of model evaluations was high, the EnKF and 4DVar methods yielded comparable performance. However, when the number of model evaluations was restricted, the errors in the EnKF assimilations were found to be considerably larger than those of the variational method. In addition, two strategies for the mitigation of data assimilation induced $\vec{\nabla} \cdot \vec{B}$ errors were also assessed here: Powell's 8-wave ideal MHD formulation and a divergence-free projection strategy for the magnetic field. In the case of the EnKF method, the projection approach was found to be more effective as compared to Powell's method, albeit incurring higher computational cost and exacerbating the aforementioned stability issues present with the EnKF. On the other hand, for the variational method, both $\vec{\nabla} \cdot \vec{B}$ error handling strategies performed quite similarly provided that the background covariance matrix is made to respect the $\vec{\nabla} \cdot \vec{B} = 0$ condition. With these strategies in place, the reliable assimilation of all components of magnetic field observations was shown to be possible, thereby overcoming a major challenge in MHD-based data assimilation [57].

In spite of the preceding observed shortcomings of sequential method, it is noted that the EnKF approach remains an effective algorithm for data assimilation with a strong potential in applications related to space weather forecasting. Unlike the 4DVar method, sequential methods are generally more readily implemented in parallel on current and future high-performance computing architectures and can provide posterior error statistics. Conversely, the most significant drawback of 4DVar approach is the considerable time investment required to develop and implement the adjoint model. Nevertheless, for the ideal MHD equations of interest in the current study, this difficulty is greatly alleviated by the derivation of the semi-analytical adjoint model presented herein. It is also noted that the proposed adjoint model may also be used apart from the data assimilation application considered here as a sensitivity analysis tool.

The numerical experiments presented herein demonstrate the potential utility of global MHD models for performing space weather forecasts, despite the incorrect model inputs. Furthermore, the results of the DA comparisons for the tree 1D initial value problems provide strong encouragement to pursue variational data assimilation for more complex simulations of MHD flows relevant to space weather phenomena. In particular, the authors are currently pursuing the application of variational data assimilation to enhance the predictions of fully 3D MHD flows. Nevertheless, some important challenges not considered here must first be addressed. In particular, an extension of the adjoint model developed herein to fully 3D simulations must be developed. To this end, Eq. (50) must be expanded to account for all 26 neighbouring cells in a typical second-order 3D finite volume stencil. While the numerical-flux Jacobians can be evaluated in an identical manner, Jacobians associated with spatial reconstruction will see considerable change compared to their 1D analogues. Additionally, further assessments of mitigation strategies for data assimilation induced $\vec{\nabla} \cdot \vec{B}$ errors, particularly in the context of 3D simulations, should be pursued. Finally, further desirable extensions of the current variational DA method

for the ideal MHD equations would involve the inclusion of model parameters beyond the initial conditions in the assimilation strategy. Examples of the latter would include DA treatments for boundary conditions and/or empirical model parameters, such as those affecting intrinsic magnetic fields.

CRediT authorship contribution statement

J.H. Arnal: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **C.P.T. Groth:** Writing – review & editing, Software, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: J. H. Arnal reports financial support was provided by Government of Ontario. C. P. T. Groth reports financial support was provided by Natural Sciences and Engineering Research Council of Canada.

Data availability

Data will be made available on request.

Acknowledgments

Appendix A. 4DVar preconditioning

Although Eq. (31) may be minimized in its original form, faster convergence may be achieved by employing a change of variables and thereby preconditioning the minimization problem. Following Courtier [31] and Lorenc [30], a new control variable given by

$$\mathbf{V} = \sqrt{(\mathbf{\Sigma}_\mathbf{U}^0)^{-1}}(\mathbf{U}^0 - \mathbf{U}^b), \tag{A.1}$$

is introduced. After writing the cost function, J, in terms of $\mathbf{V}$ instead of $\mathbf{U}^0$, it may be shown that the gradient of J with respect to $\mathbf{V}$ is given by

$$\frac{\partial J}{\partial \mathbf{V}} = \mathbf{V} - \sqrt{\mathbf{\Sigma}_\mathbf{U}^0}\lambda^0. \tag{A.2}$$

Minimizing with respect to $\mathbf{V}$ speeds convergence in modes where the background, $\mathbf{U}^b$, is inaccurate [30]. In addition, employing the change of variables removes the need to invert $\mathbf{\Sigma}_{\mathbf{U}^0\mathbf{U}^0}$. However, its square-root form is now required.

Appendix B. Expressions in adjoint model

B.1. Differentiation of the RK2 scheme

By writing the fully discrete ideal MHD equations in terms of the RK2 stages,

$$\mathcal{M}(\mathbf{U}^n) = \mathbf{U}^{n+1}(\mathbf{U}^n, \widetilde{\mathbf{U}}^{n+1}(\mathbf{U}^n)), \tag{B.1}$$

and applying the chain-rule, one can obtain

$$\frac{\partial \mathcal{M}(\mathbf{U}^n)}{\partial \mathbf{U}^n} = \frac{\partial \mathbf{U}^{n+1}}{\partial \mathbf{U}^n} + \frac{\partial \mathbf{U}^{n+1}}{\partial \widetilde{\mathbf{U}}^{n+1}} \frac{\partial \widetilde{\mathbf{U}}^{n+1}}{\partial \mathbf{U}^n}, \tag{B.2}$$

$$= \frac{1}{2}\mathbf{I} + \frac{1}{2}\left[\mathbf{I} + \Delta t \frac{\partial \mathbf{R}(\widetilde{\mathbf{U}}^{n+1})}{\partial \widetilde{\mathbf{U}}^{n+1}}\right]\left[\mathbf{I} + \Delta t \frac{\partial \mathbf{R}(\mathbf{U}^n)}{\partial \mathbf{U}^n}\right]. \tag{B.3}$$

A two-stage representation can then be obtained by defining

$$\widetilde{\lambda}^n = \left[\mathbf{I} + \Delta t \frac{\partial \mathbf{R}(\widetilde{\mathbf{U}}^{n+1})}{\partial \widetilde{\mathbf{U}}^{n+1}}\right]^{\mathsf{T}} \lambda^{n+1}, \tag{B.4}$$

$$= \lambda^{n+1} + \Delta t \left[\frac{\partial \mathbf{R}(\widetilde{\mathbf{U}}^{n+1})}{\partial \widetilde{\mathbf{U}}^{n+1}}\right]^{\mathsf{T}} \lambda^{n+1}, \tag{B.5}$$

which recovers Eq. (45). Multiplying the transpose of Eq. (B.3) by λ^{n+1}, and substituting in Eq. (B.4) results in

$$\left[\frac{\partial \mathcal{M}}{\partial \mathbf{U}^n}\right]^{\mathsf{T}} \lambda^{n+1} = \frac{1}{2}\left(\lambda^{n+1} + \widetilde{\lambda}^n + \Delta t \left[\frac{\partial \mathbf{R}(\mathbf{U}^n)}{\partial \mathbf{U}^n}\right]^{\mathsf{T}} \widetilde{\lambda}^n\right), \tag{B.6}$$

which recovers Eq. (46).

B.2. Jacobian matrices of the adjoint model

Instead of working with Eq. (49), its transpose is considered, i.e.,

$$\left[\frac{\partial \mathbf{R}^n}{\partial \mathbf{U}^n}\right]^{\mathsf{T}} = \left[\frac{\partial \mathbf{W}^n}{\partial \mathbf{U}^n}\right]^{\mathsf{T}} \left(\left[\frac{\partial \mathbf{W}_L^n}{\partial \mathbf{W}^n}\right]^{\mathsf{T}} \left[\frac{\partial \boldsymbol{\mathcal{F}}^n}{\partial \mathbf{W}_L^n}\right]^{\mathsf{T}} + \left[\frac{\partial \mathbf{W}_R^n}{\partial \mathbf{W}^n}\right]^{\mathsf{T}} \left[\frac{\partial \boldsymbol{\mathcal{F}}^n}{\partial \mathbf{W}_R^n}\right]^{\mathsf{T}}\right) \left[\frac{\partial \mathbf{R}^n}{\partial \boldsymbol{\mathcal{F}}^n}\right]^{\mathsf{T}}. \tag{B.7}$$

Each of the above matrices are sparse block matrices with 8×8 sized elements corresponding to the 8 conserved variables.

The Jacobian transpose, $[\partial \mathbf{W}^n / \partial \mathbf{U}^n]^{\mathsf{T}}$, is a diagonal matrix of size $M \times M$ with diagonal elements given by

$$\left[\frac{\partial \mathbf{W}^n}{\partial \mathbf{U}^n}\right]_{ii}^{\mathsf{T}} = \left[\frac{\partial \mathbf{W}_i^n}{\partial \mathbf{U}_i^n}\right]^{\mathsf{T}} = \begin{bmatrix} 1 & -u_x/\rho & -u_y/\rho & -u_z/\rho & 0 & 0 & 0 & -g\vec{u}\cdot\vec{u}/2 \\ 0 & 1/\rho & 0 & 0 & 0 & 0 & 0 & gu_x \\ 0 & 0 & 1/\rho & 0 & 0 & 0 & 0 & gu_y \\ 0 & 0 & 0 & 1/\rho & 0 & 0 & 0 & gu_z \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & gB_x \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & gB_y \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & gB_z \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -g \end{bmatrix}, \tag{B.8}$$

where $g = 1 - \gamma$ and variables are evaluated at the ith cell. Similarly, $[\partial \mathbf{R}^n / \partial \boldsymbol{\mathcal{F}}^n]^{\mathsf{T}}$ is a matrix of size $M + 1 \times M$ given by

$$\left[\frac{\partial \mathbf{R}^n}{\partial \boldsymbol{\mathcal{F}}^n}\right]^{\mathsf{T}} = \frac{1}{\Delta x}\begin{bmatrix} \mathbf{I} & & & \\ -\mathbf{I} & \mathbf{I} & & \\ & -\mathbf{I} & \ddots & \\ & & \ddots & \mathbf{I} \\ & & & -\mathbf{I} \end{bmatrix}. \tag{B.9}$$

Note that the product of $[\partial \mathbf{R}^n / \partial \boldsymbol{\mathcal{F}}^n]^{\mathsf{T}}$ with an adjoint vector, λ, is equivalent to taking the backwards difference of the elements of λ, i.e.,

$$\left[\frac{\partial \mathbf{R}^n}{\partial \boldsymbol{\mathcal{F}}^n}\right]^{\mathsf{T}} \lambda = [\lambda_1, \nabla\lambda_2, \nabla\lambda_3, \dots, \nabla\lambda_M, -\lambda_M]^{\mathsf{T}}. \tag{B.10}$$

This property is used in the derivation of Eq. (50).

Finally, $[\partial \mathbf{W}_L^n / \partial \mathbf{W}^n]^{\mathsf{T}}$ and $[\partial \mathbf{W}_R^n / \partial \mathbf{W}^n]^{\mathsf{T}}$ are each of size $M \times M + 1$. These matrices are block tridiagonal with an extra block column, and

are given in Box I It is easy to show that the sub-matrices may be written in terms of derivatives of the limiters and solution gradient according to

$$\left[\frac{\partial \mathbf{W}_{L,i-1/2}}{\partial \mathbf{W}_i}\right]^{\mathsf{T}} = \left[\frac{\partial \boldsymbol{\phi}_{i-1}}{\partial \mathbf{W}_i}\right]^{\mathsf{T}} \circ \frac{\partial \mathbf{W}_{i-1}}{\partial x}\frac{\Delta x}{2} + \boldsymbol{\phi}_{i-1} \circ \left[\frac{\partial}{\partial \mathbf{W}_i}\left(\frac{\partial \mathbf{W}_{i-1}}{\partial x}\right)\right]^{\mathsf{T}}\frac{\Delta x}{2} \tag{B.13}$$

$$\left[\frac{\partial \mathbf{W}_{L,i+1/2}}{\partial \mathbf{W}_i}\right]^{\mathsf{T}} = \mathbf{I} + \left[\frac{\partial \boldsymbol{\phi}_i}{\partial \mathbf{W}_i}\right]^{\mathsf{T}} \circ \frac{\partial \mathbf{W}_i}{\partial x}\frac{\Delta x}{2} + \boldsymbol{\phi}_i \circ \left[\frac{\partial}{\partial \mathbf{W}_i}\left(\frac{\partial \mathbf{W}_i}{\partial x}\right)\right]^{\mathsf{T}}\frac{\Delta x}{2} \tag{B.14}$$

$$\left[\frac{\partial \mathbf{W}_{L,i+3/2}}{\partial \mathbf{W}_i}\right]^{\mathsf{T}} = \left[\frac{\partial \boldsymbol{\phi}_{i+1}}{\partial \mathbf{W}_i}\right]^{\mathsf{T}} \circ \frac{\partial \mathbf{W}_{i+1}}{\partial x}\frac{\Delta x}{2} + \boldsymbol{\phi}_{i+1} \circ \left[\frac{\partial}{\partial \mathbf{W}_i}\left(\frac{\partial \mathbf{W}_{i+1}}{\partial x}\right)\right]^{\mathsf{T}}\frac{\Delta x}{2} \tag{B.15}$$

$$\left[\frac{\partial \mathbf{W}_{R,i-3/2}}{\partial \mathbf{W}_i}\right]^{\mathsf{T}} = -\left[\frac{\partial \mathbf{W}_{L,i-1/2}}{\partial \mathbf{W}_i}\right]^{\mathsf{T}}, \tag{B.16}$$

$$\left[\frac{\partial \mathbf{W}_{R,i-1/2}}{\partial \mathbf{W}_i}\right]^{\mathsf{T}} = 2\mathbf{I} - \left[\frac{\partial \mathbf{W}_{L,i+1/2}}{\partial \mathbf{W}_i}\right]^{\mathsf{T}}, \tag{B.17}$$

$$\left[\frac{\partial \mathbf{W}_{R,i+1/2}}{\partial \mathbf{W}_i}\right]^{\mathsf{T}} = -\left[\frac{\partial \mathbf{W}_{L,i+3/2}}{\partial \mathbf{W}_i}\right]^{\mathsf{T}}. \tag{B.18}$$

The derivatives of the solution gradient appearing in the expressions above are obtained by differentiating Eq. (8), resulting in

$$\left[\frac{\partial}{\partial \mathbf{W}_i}\left(\frac{\partial \mathbf{W}_{i-1}}{\partial x}\right)\right]^{\mathsf{T}} = \frac{\Delta x_{i-1}}{(\Delta x_{i-1})^2 + (\nabla x_{i-1})^2}\mathbf{I},$$
$$\left[\frac{\partial}{\partial \mathbf{W}_i}\left(\frac{\partial \mathbf{W}_{i+1}}{\partial x}\right)\right]^{\mathsf{T}} = -\frac{\nabla x_{i+1}}{(\Delta x_{i+1})^2 + (\nabla x_{i+1})^2}\mathbf{I}, \tag{B.19}$$

$$\left[\frac{\partial}{\partial \mathbf{W}_i}\left(\frac{\partial \mathbf{W}_i}{\partial x}\right)\right]^{\mathsf{T}} = \frac{\nabla x_i - \Delta x_i}{(\Delta x_i)^2 + (\nabla x_i)^2}\mathbf{I}. \tag{B.20}$$

The additional derivatives of the limiter, $\boldsymbol{\phi}_i$, are more involved to evaluate. As $\boldsymbol{\phi}_i$ does not involve operations between conserved variables, $\boldsymbol{\phi}_i$ and its derivatives may be expressed in terms of the arbitrary variable, w_i. In this way, the Jacobians, $[\partial \boldsymbol{\phi}_j / \partial \mathbf{W}_i]^{\mathsf{T}}$, found in Eqs. (B.13)–(B.18), may be expressed in a diagonal form given in Box II. To proceed, the limiter is written in its general form

$$\phi_i(w_i) = \min(\phi_{i+1/2}, \phi_{i-1/2}), \tag{B.22}$$

where

$$\phi_{i\pm1/2} = \begin{cases} \varphi(\hat{r}_{i\pm1/2}) & \text{if} \quad w_{i\pm1/2} - w_i > 0 \\ \varphi(\check{r}_{i\pm1/2}) & \text{if} \quad w_{i\pm1/2} - w_i < 0 \\ 1 & \text{if} \quad w_{i\pm1/2} - w_i = 0 \end{cases}, \qquad w_{i\pm1/2} = w_i \pm \frac{\partial w_i}{\partial x}\frac{\Delta x}{2}, \tag{B.23}$$

$$\hat{r}_{i\pm1/2} = \frac{w_{\max} - w_i}{w_{i\pm1/2} - w_i} \qquad \check{r}_{i\pm1/2} = \frac{w_{\min} - w_i}{w_{i\pm1/2} - w_i}, \tag{B.24}$$

$$w_{\max} = \max(w_{i-1}, w_i, w_{i+1}), \qquad w_{\min} = \min(w_{i-1}, w_i, w_{i+1}). \tag{B.25}$$

Here, φ is the limiter function of choice. Differentiating, derivatives of $\phi_i(w_i)$ may now be written as

$$\frac{\partial \phi_j}{\partial w_i} = \begin{cases} \frac{\partial \phi_{j+1/2}}{\partial w_i} & \text{if} \quad \phi_{j+1/2} < \phi_{j-1/2} \\ \frac{\partial \phi_{j-1/2}}{\partial w_i} & \text{if} \quad \phi_{j-1/2} < \phi_{j+1/2} \end{cases}, \tag{B.26}$$

where

$$\frac{\partial \phi_{j\pm1/2}}{\partial w_i} = \begin{cases} \frac{d\varphi(\hat{r}_{j\pm1/2})}{d\hat{r}_{j\pm1/2}}\frac{\partial \hat{r}_{j\pm1/2}}{\partial w_i} & \text{if} \quad w_{j\pm1/2} - w_j > 0 \\ \frac{d\varphi(\check{r}_{j\pm1/2})}{d\check{r}_{j\pm1/2}}\frac{\partial \check{r}_{j\pm1/2}}{\partial w_i} & \text{if} \quad w_{j\pm1/2} - w_j < 0 \\ 0 & \text{if} \quad w_{j\pm1/2} - w_j = 0 \end{cases} \tag{B.27}$$

$$\left[\frac{\partial \mathbf{W}_L^n}{\partial \mathbf{W}^n}\right]^\mathsf{T} = \begin{bmatrix} \left[\frac{\partial \mathbf{W}_{L,1/2}}{\partial \mathbf{W}_1}\right]^\mathsf{T} & \left[\frac{\partial \mathbf{W}_{L,1+1/2}}{\partial \mathbf{W}_1}\right]^\mathsf{T} & \left[\frac{\partial \mathbf{W}_{L,1+3/2}}{\partial \mathbf{W}_1}\right]^\mathsf{T} \\ & \ddots & \ddots & \ddots \\ & & \left[\frac{\partial \mathbf{W}_{L,i-1/2}}{\partial \mathbf{W}_i}\right]^\mathsf{T} & \left[\frac{\partial \mathbf{W}_{L,i+1/2}}{\partial \mathbf{W}_i}\right]^\mathsf{T} & \left[\frac{\partial \mathbf{W}_{L,i+3/2}}{\partial \mathbf{W}_i}\right]^\mathsf{T} \\ & & & \ddots & \ddots & \ddots \\ & & & & \left[\frac{\partial \mathbf{W}_{L,M-1/2}}{\partial \mathbf{W}_M}\right]^\mathsf{T} & \left[\frac{\partial \mathbf{W}_{L,M+1/2}}{\partial \mathbf{W}_M}\right]^\mathsf{T} \end{bmatrix}, \tag{B.11}$$

$$\left[\frac{\partial \mathbf{W}_R^n}{\partial \mathbf{W}^n}\right]^\mathsf{T} = \begin{bmatrix} \left[\frac{\partial \mathbf{W}_{R,1/2}}{\partial \mathbf{W}_1}\right]^\mathsf{T} & \left[\frac{\partial \mathbf{W}_{R,1+1/2}}{\partial \mathbf{W}_1}\right]^\mathsf{T} \\ \ddots & \ddots \\ & \left[\frac{\partial \mathbf{W}_{R,i-3/2}}{\partial \mathbf{W}_i}\right]^\mathsf{T} & \left[\frac{\partial \mathbf{W}_{R,i-1/2}}{\partial \mathbf{W}_i}\right]^\mathsf{T} & \left[\frac{\partial \mathbf{W}_{R,i+1/2}}{\partial \mathbf{W}_i}\right]^\mathsf{T} \\ & & \ddots & \ddots & \ddots \\ & & \left[\frac{\partial \mathbf{W}_{R,M-3/2}}{\partial \mathbf{W}_M}\right]^\mathsf{T} & \left[\frac{\partial \mathbf{W}_{R,M-1/2}}{\partial \mathbf{W}_M}\right]^\mathsf{T} & \left[\frac{\partial \mathbf{W}_{R,M+1/2}}{\partial \mathbf{W}_M}\right]^\mathsf{T} \end{bmatrix}, \tag{B.12}$$

Box I.

$$\left[\frac{\partial \boldsymbol{\phi}_j(\mathbf{W}_j)}{\partial \mathbf{W}_i}\right]^\mathsf{T} = \begin{bmatrix} \frac{\partial \phi_i(\rho)}{\partial \rho} \\ & \frac{\partial \phi_i(u_x)}{\partial u_x} \\ & & \frac{\partial \phi_i(u_y)}{\partial u_y} \\ & & & \frac{\partial \phi_i(u_z)}{\partial u_z} \\ & & & & \frac{\partial \phi_i(B_x)}{\partial B_x} \\ & & & & & \frac{\partial \phi_i(B_y)}{\partial B_y} \\ & & & & & & \frac{\partial \phi_i(B_z)}{\partial B_z} \\ & & & & & & & \frac{\partial \phi_i(p)}{\partial p} \end{bmatrix} \tag{B.21}$$

Box II.

$$\frac{\partial \hat{r}_{j\pm1/2}}{\partial w_i} = \frac{\pm\left(\frac{\partial w_{\max}}{\partial w_i} - \delta_{ij}\right)\frac{\partial w_j}{\partial x}\frac{\Delta x}{2} \mp (w_{\max} - w_j)\frac{\partial}{\partial w_i}\left(\frac{\partial w_j}{\partial x}\right)\frac{\Delta x}{2}}{\left(\frac{\partial w_j}{\partial x}\frac{\Delta x}{2}\right)^2}, \tag{B.28}$$

$$\frac{\partial \check{r}_{j\pm1/2}}{\partial w_i} = \frac{\pm\left(\frac{\partial w_{\min}}{\partial w_i} - \delta_{ij}\right)\frac{\partial w_j}{\partial x}\frac{\Delta x}{2} \mp (w_{\min} - w_j)\frac{\partial}{\partial w_i}\left(\frac{\partial w_j}{\partial x}\right)\frac{\Delta x}{2}}{\left(\frac{\partial w_j}{\partial x}\frac{\Delta x}{2}\right)^2}, \tag{B.29}$$

$$\frac{\partial w_{\max}}{\partial w_i} = \begin{cases} 1 & \text{if } w_{\max} = w_i \\ 0 & \text{otherwise} \end{cases}, \qquad \frac{\partial w_{\min}}{\partial w_i} = \begin{cases} 1 & \text{if } w_{\min} = w_i \\ 0 & \text{otherwise} \end{cases}, \tag{B.30}$$

and δ_{ij} is the Kronecker delta function. The limiter function, φ, and its derivative are left unspecified in the above equations. In the present work, the Vanleer Leer limiter function was used [74] which, along with its derivative, may be written as

$$\varphi(r) = \frac{2r}{1+r}, \qquad \frac{d\varphi(r)}{dr} = \frac{2}{(1+r)^2}. \tag{B.31}$$

The use of other limiter functions only requires swapping Eq. (B.31) by the appropriate relations.

Appendix C. Jacobian matrices associated with the Powell source term

$$\left[\frac{\partial \mathbf{S}_{i\pm1}}{\partial \mathbf{U}_i}\right]^\mathsf{T} = \pm\frac{1}{2\Delta x} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & B_x & B_y & B_z & u_x & u_y & u_z & \vec{u}\cdot\vec{B} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{i\pm1}, \tag{C.1}$$

$$\left[\frac{\partial \mathbf{S}_i}{\partial \mathbf{U}_i}\right]^\mathsf{T} = -\frac{Bx_{i+1} - Bx_{i-1}}{2\Delta x} \begin{bmatrix} 0 & 0 & 0 & 0 & -\frac{u_x}{\rho} & -\frac{u_y}{\rho} & -\frac{u_z}{\rho} & -\frac{\vec{u}\cdot\vec{B}}{\rho} \\ 0 & 0 & 0 & 0 & \frac{1}{\rho} & 0 & 0 & \frac{B_x}{\rho} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{\rho} & 0 & \frac{B_y}{\rho} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\rho} & \frac{B_x}{\rho} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & u_x \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & u_y \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & u_z \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}_i. \tag{C.2}$$

References

[1] Kane R. The idea of space weather — A historical perspective. Adv Space Res 2006;37(6):1261–4.

[2] Schrijver CJ, Kauristie K, Aylward AD, Denardini CM, Gibson SE, Glover A, et al. Understanding space weather to shield society: A global road

map for 2015–2025 commissioned by COSPAR and ILWS. Adv Space Res 2015;55(12):2745–807. http://dx.doi.org/10.1016/j.asr.2015.03.023.

[3] Warner TT. Numerical Weather and Climate Prediction. Cambridge: Cambridge University Press; 2010.

[4] Powell KG, Roe PL, Linde TJ, Gombosi TI, De Zeeuw DL. A solution-adaptive upwind scheme for ideal magnetohydrodynamics. J Comput Phys 1999;154:284–309.

[5] Groth CPT, De Zeeuw DL, Gombosi TI, Powell KG. Global three-dimensional MHD simulation of a space weather event: CME formation, interplanetary propagation, and interaction with the magnetosphere. J Geophys Res 2000;105(A11):25,053–78.

[6] Roussev II, Gombosi TI, Sokolov IV, Velli M, Manchester W, DeZeeuw DL, et al. A three-dimensional model of the solar wind incorporating solar magnetogram observations. Astrophys J Lett 2003;595(1):L57. http://dx.doi.org/10.1086/378878.

[7] Manchester WB, Gombosi TI, Roussev I, De Zeeuw DL, Sokolov IV, Powell KG, et al. Three-dimensional MHD simulation of a flux rope driven CME. J Geophys Res Space Phys 2004;109(A1):A01102. http://dx.doi.org/10.1029/2002JA009672.

[8] Tóth G, Sokolov IV, Gombosi TI, Chesney DR, Clauer CR, De Zeeuw DL, et al. Space weather modeling framework: A new tool for the space science community. J Geophys Res Space Phys 2005;110(A12):A12226. http://dx.doi.org/10.1029/2005JA011126.

[9] Tóth G, De Zeeuw DL, Gombosi TI, Manchester WB, Ridley AJ, Sokolov IV, et al. Sun-to-thermosphere simulation of the 28-30 October 2003 storm with the space weather modeling framework. Space Weather 2007;5(6):S06003–1––S06003–16. http://dx.doi.org/10.1029/2006SW000272.

[10] Manchester WB, Vourlidas A, Tóth G, Lugaz N, Roussev II, Sokolov IV, et al. Three-dimensional MHD simulation of the 2003 October 28 coronal mass ejection: Comparison with LASCO coronagraph observations. Astrophys J 2008;684(2):1448–60. http://dx.doi.org/10.1086/590231.

[11] Tóth G, van der Holst B, Sokolov IV, De Zeeuw DL, Gombosi TI, Fang F, et al. Adaptive numerical algorithms in space weather modeling. J Comput Phys 2012;231(3):870–903. http://dx.doi.org/10.1016/j.jcp.2011.02.006.

[12] van der Holst B, Sokolov IV, Meng X, Jin M, Manchester WB, Tóth G, et al. Alfvén wave solar model (AWSoM): Coronal heating. Astrophys J 2014;782(2):81–95. http://dx.doi.org/10.1088/0004-637X/782/2/81.

[13] Odstrčil D, Riley P, Zhao XP. Numerical simulation of the 12 May 1997 interplanetary CME event. J Geophys Res Space Phys 2004;109(A2):A02116. http://dx.doi.org/10.1029/2003JA010135.

[14] Odstrčil D, Pizzo VJ, Linker JA, Riley P, Lionello R, Mikić Z. Initial coupling of coronal and heliospheric numerical magnetohydrodynamic codes. J Atm Sol-Terr Phys 2004;66(15):1311–20. http://dx.doi.org/10.1016/j.jastp.2004.04.007.

[15] Odstrčil D, Pizzo VJ, Arge CN, Bissi MM, Hick PP, Jackson BV, et al. Numerical simulations of solar wind disturbances by coupled models. In: Pogorelov NV, Audit E, Zank GP, editors. Numerical modeling of space plasma flows. Astronomical society of the Pacific conference series, vol. 385, 2008, p. 167.

[16] Pomoell J, Poedts S. EUHFORIA: European heliospheric forecasting information asset. J Space Weather Space Clim 2018;8(A35). http://dx.doi.org/10.1051/swsc/2018020.

[17] Poedts S, Lani A, Scolini C, Verbeke C, Wijsen N, Lapenta G, et al. EUropean heliospheric FORecasting information asset 2.0. J Space Weather Space Clim 2020;10(A57). http://dx.doi.org/10.1051/swsc/2020055.

[18] Narechania NM, Nikolić L, Freret L, De Sterck H, Groth CPT. An integrated data-driven solar wind — CME numerical framework for space weather forecasting. J Space Weather Space Clim 2021;11(8). http://dx.doi.org/10.1051/swsc/2020068.

[19] Kalnay E. Atmospheric modeling, data assimilation and predictability. Cambridge: Cambridge University Press; 2003.

[20] Park SK, Xu L. Data assimilation for atmospheric, oceanic and hydrologic applications, vol. II, New York: Springer; 2013.

[21] Evensen G, Vossepoel FC, van Leeuwen PJ. Data assimilation fundamentals: A unified formulation of the state and parameter estimation problem. Springer; 2022.

[22] Murray SA. The importance of ensemble techniques for operational space weather forecasting. Space Weather 2018;16(7):777–83. http://dx.doi.org/10.1029/2018SW001861.

[23] Morley SK, Welling DT, Woodroffe JR. Perturbed input ensemble modeling with the space weather modeling framework. Space Weather 2018;16(9):1330–47. http://dx.doi.org/10.1029/2018SW002000.

[24] Kalman RE. A new approach to linear filtering and prediction problems. ASME J Basic Eng 1960;82:35–45.

[25] Evensen G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J Geophys Res: Oceans 1994;99(C5):10143–62.

[26] Burgers G, van Leeuwen PJ, Evensen G. Analysis scheme in the ensemble Kalman filter. Mon Weather Rev 1998;126(7):1719–24.

[27] Whitaker JS, Hamill TM. Ensemble data assimilation without perturbed observations. Mon Weather Rev 2002;130(7):1913–24.

[28] Hunt BR, Kostelich EJ, Szunyogh I. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. Physica D 2007;230(1–2):112–26.

[29] Le Dimet F-X, Talagrand O. Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. Tellus A 1986;38(2):97–110.

[30] Lorenc AC. Optimal nonlinear objective analysis. Q J R Meteorol Soc 1988;114(479):205–40.

[31] Courtier P. Dual formulation of four-dimensional variational assimilation. Q J R Meteorol Soc 1997;123(544):2449–61.

[32] Mons V, Chassaing J-C, Gomez T, Sagaut P. Reconstruction of unsteady viscous flows using data assimilation schemes. J Comput Phys 2016;316:255–80.

[33] Lang M, Owens MJ. A variational approach to data assimilation in the solar wind. Space Weather 2019;17(1):59–83. http://dx.doi.org/10.1029/2018SW001857.

[34] Carrassi A, Bocquet M, Bertino L, Evensen G. Data assimilation in the geosciences: An overview of methods, issues, and perspectives. Wiley Interdisciplinary Reviews: Climate Change 2018;9(5):e535.

[35] Lorenc AC. The potential of the ensemble Kalman filter for NWP–A comparison with 4D-var. Q J R Meteorol Soc 2003;129(595):3183–203.

[36] Caya A, Sun J, Snyder C. A comparison between the 4DVAR and the ensemble Kalman filter techniques for radar data assimilation. Mon Weather Rev 2005;133(11):3081–94.

[37] Kalnay E, Li H, Miyoshi T, Shu-Chih Y, Ballabrera-Poy J. 4-D-var or ensemble Kalman filter? Tellus A 2007;59(5):758–73.

[38] Miyoshi T, Sato Y, Kadowaki T. Ensemble Kalman filter and 4D-Var intercomparison with the Japanese operational global analysis and prediction system. Mon Weather Rev 2010;138(7):2846–66.

[39] Skachko S, Ménard R, Errera Q, Christophe Y, Chabrillat S. EnKF and 4D-Var data assimilation with chemical transport model BASCOE (version 05.06). Geosci Model Dev 2016;9(8):2893–908.

[40] Liu J, Bowman KW, Lee M. Comparison between the local ensemble transform Kalman filter (LETKF) and 4D-Var in atmospheric CO2 flux inversion with the goddard earth observing system-chem model and the observation impact diagnostics from the LETKF. J Geophys Res: Atmos 2016;121(21):13–066.

[41] Buehner M, Houtekamer PL, Charette C, Mitchell HL, He B. Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. Mon Weather Rev 2010;138(5):1567–86.

[42] Zhang M, Zhang F, Huang X, Zhang X. Intercomparison of an ensemble Kalman filter with three-and four-dimensional variational data assimilation methods in a limited-area model over the month of June 2003. Mon Weather Rev 2011;139(2):566–72.

[43] Chatterjee A, Michalak AM. Comparison of ensemble Kalman filter and variational approaches for CO2 data assimilation. Atmos Chem Phys 2013;13(23):11643–60.

[44] Fairbairn D, Pring SR, Lorenc AC, Roulstone I. A comparison of 4DVar with ensemble data assimilation methods. Q J R Meteorol Soc 2014;140(678):281–94.

[45] Kabir SKF, Assumaning GA, Chang S. Efficiency of using 4DVar, 3DVar and EnKF data assimilation methods in groundwater contaminant transport modelling. Eur J Environ Civ Eng 2019;23(4):515–31.

[46] Gopalakrishnan G, Hoteit I, Cornuelle BD, Rudnick DL. Comparison of 4DVAR and EnKF state estimates and forecasts in the gulf of Mexico. Q J R Meteorol Soc 2019;145(721):1354–76.

[47] Sun Z, Tangborn A, Kuang W. Data assimilation in a sparsely observed one-dimensional modeled MHD system. Nonlinear Process Geophys 2007;14(2):181–92.

[48] Fournier A, Eymin C, Alboussière T. A case for variational geomagnetic data assimilation: Insights from a one-dimensional, nonlinear, and sparsely observed MHD system. Nonlinear Process Geophys 2007;14(2):163–80.

[49] Ren Z, Guo S, Li Z, Wu Z. Adjoint-based parameter and state estimation in 1-D magnetohydrodynamic (MHD) flow system. J Ind Manag Optim 2018;14(4):1579–94.

[50] Mendoza OB, De Moor B, Bernstein DS. Data assimilation for magnetohydrodynamics systems. J Comput Appl Math 2006;189(1–2):242–59.

[51] Teixeira BOS, Ridley A, Tôrres LAB, Aguirre LA, Bernstein DS. Data assimilation for magnetohydrodynamics with a zero-divergence constraint on the magnetic field. In: 2008 American control conference. IEEE; 2008, p. 2534–9.

[52] Biswas A, Hudson J, Larios A, Pei Y. Continuous data assimilation for the 2D magnetohydrodynamic equations using one component of the velocity and magnetic fields. Asymptot Anal 2018;108(1–2):1–43.

[53] Hudson J, Jolly M. Numerical efficacy study of data assimilation for the 2D magnetohydrodynamic equations. J Comput Dyn 2019;6(1):131–45.

[54] Merkin VG, Kondrashov D, Ghil M, Anderson BJ. Data assimilation of low-altitude magnetic perturbations into a global magnetosphere model. Space Weather 2016;14(2):165–84. http://dx.doi.org/10.1002/2015SW001330.

[55] Riley P, Lionello R. Mapping solar wind streams from the Sun to 1 AU: A comparison of techniques. Sol Phys 2011;270(2):575–92.

[56] Li K, Jackson A, Livermore PW. Variational data assimilation for a forced, inertia-free magnetohydrodynamic dynamo model. Geophys J Int 2014;199(3):1662–76.

[57] Lang M, Browne P, van Leeuwen PJ, Owens M. Data assimilation in the solar wind: Challenges and first results. Space Weather 2017;15:1490–510. http://dx.doi.org/10.1002/2017SW001681.

[58] Jivani A, Sachdeva N, Huang Z, Chen Y, van der Holst B, Manchester W, et al. Global sensitivity analysis and uncertainty quantification for background solar wind using the Alfvén wave solar atmosphere model. Space Weather 2023;21(1):e2022SW003262. http://dx.doi.org/10.1029/2022SW003262.

[59] Arge CN, Henney CJ, Koller J, Compeau CR, Young S, MacKenzie D, et al. Air force data assimilative photospheric flux transport (ADAPT) model. AIP Conf Proc 2010;1216:343–436. http://dx.doi.org/10.1063/1.3395870.

[60] Hickmann KS, Godinez HC, Henney CJ, Arge CN. Data assimilation in the ADAPT photospheric flux transport model. Sol Phys 2015;290(4):1105–18.

[61] Innocenti ME, Lapenta G, Vršnak B, Crespon F, Skandrani C, Temmer M, et al. Improved forecasts of solar wind parameters using the Kalman filter. Space Weather 2011;9(10):1–15. http://dx.doi.org/10.1029/2011SW000659.

[62] Meadors GD, Jones SI, Hickmann KS, Arge CN, Godinez-Vasquez HC, Henney CJ. Data assimilative optimization of WSA source surface and interface radii using particle filtering. Space Weather 2020;18(5):1–17. http://dx.doi.org/10.1029/2020SW002464.

[63] Wang YM, Sheeley NR. Solar wind speed and coronal flux-tube expansion. Astrophys J 1990;355:726–32.

[64] Arge CN, Pizzo VJ. Improvement in the prediction of solar wind conditions using near-real time solar magnetic field updates. J Geophys Res Space Phys 2000;105(A5):10,465–79. http://dx.doi.org/10.1029/1999JA000262.

[65] Gombosi TI. Physics of the space environment. Cambridge: Cambridge University Press; 1998.

[66] Godunov SK. Finite-difference method for numerical computations of discontinuous solutions of the equations of fluid dynamics. Mat Sb 1959;47:271–306.

[67] Ivan L, De Sterck H, Northrup SA, Groth CPT. Three-dimensional MHD on cubed-sphere grids: Parallel solution-adaptive simulation framework. Paper 2011-3382, AIAA; 2011.

[68] Ivan L, De Sterck H, Northrup SA, Groth CPT. Hyperbolic conservation laws on three-dimensional cubed-sphere grids: A parallel solution-adaptive simulation framework. J Comput Phys 2013;255:205–27.

[69] Susanto A, Ivan L, De Sterck H, Groth CPT. High-order central ENO finite-volume scheme for ideal MHD. J Comput Phys 2013;250:141–64.

[70] Ivan L, H. D, Susanto A, Groth CPT. High-order central ENO finite-volume scheme for hyperbolic conservation laws on three-dimensional cubed-sphere grids. J Comput Phys 2015;282:157–82.

[71] Freret L, Ivan L, De Sterck H, Groth CPT. High-order finite-volume method with block-based AMR for magnetohydrodynamics flows. J Sci Comput 2019;79(1):176–208.

[72] Nguyen TB, De Sterck H, Freret L, Groth CPT. High-order implicit time-stepping with high-order CENO methods for unsteady three-dimensional CFD simulations. Internat J Numer Methods Fluids 2022;94:121–51.

[73] Freret L, Ngigi CN, Nguyen TB, De Sterck H, Groth CPT. High-order CENO finite-volume scheme with anisotropic adaptive mesh refinement: Efficient inexact Newton method for steady three-dimensional flows. J Comput Phys 2023;94(48). http://dx.doi.org/10.1007/s10915-022-02068-3.

[74] van Leer B. Towards the ultimate conservative difference scheme. II. Monotonicity and conservation combined in a second order scheme. J Comput Phys 1974;14:361–70.

[75] van Leer B. Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method. J Comput Phys 1979;32:101–36.

[76] Powell KG. An Approximate Riemann solver for magnetohydrodynamics (that works in more than one dimension). Report 94-24, ICASE; 1994.

[77] Shu C-W, Osher S. Efficient implementation of essentially non-oscillatory shock-capturing schemes. J Comput Phys 1988;77:439–71.

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

[79] Courant R, Friedrichs KO. Supersonic flow and shock waves. 5th ed. New York: Springer; 1999.

[80] Van Leeuwen PJ, Evensen G. Data assimilation and inverse methods in terms of a probabilistic formulation. Mon Weather Rev 1996;124(12):2898–913.

[81] Evensen G, van Leeuwen PJ. An ensemble Kalman smoother for nonlinear dynamics. Mon Weather Rev 2000;128(6):1852–67.

[82] Liu DC, Nocedal J. On the limited memory BFGS method for large scale optimization. Math Program 1989;45(1–3):503–28.

[83] Martins JRRA, Sturdza P, Alonso JJ. The complex-step derivative approximation. ACM Trans Math Software 2003;29(3):245–62.

[84] Marta AC, Mader CA, Martins JRRA, Van der Weide E, Alonso JJ. A methodology for the development of discrete adjoint solvers using automatic differentiation tools. Int J Comput Fluid Dyn 2007;21(9–10):307–27. http://dx.doi.org/10.1080/10618560701678647.

[85] Hogan RJ. Fast reverse-mode automatic differentiation using expression templates in C++. ACM Trans Math Software 2014;40(4):1–16.

[86] Pironneau O. On optimum design in fluid mechanics. J Fluid Mech 1974;64(1):97–110. http://dx.doi.org/10.1017/S0022112074002023.

[87] Jameson A. Aerodynamic design via control theory. J Sci Comput 1988;3(3):233–60. http://dx.doi.org/10.1007/BF01061285.

[88] Nemec M, Zingg DW. Newton-Krylov algorithm for aerodynamic design using the Navier-Stokes equations. AIAA J 2002;40(6):1146–54.

[89] Nemec M, Zingg DW. Multipoint and multi-objective aerodynamic shape optimization. AIAA J 2004;42(6):1057–65.

[90] Truong AH, Oldfield CA, Zingg DW. Mesh movement for a discrete-adjoint Newton-Krylov algorithm for aerodynamic optimization. AIAA J 2008;46(7):1695–704.

[91] Hicken JE, Zingg DW. Aerodynamic optimization algorithm with integrated geometry parameterization and mesh movement. AIAA J 2010;48(2):400–13.

[92] Becker R, Rannacher R. A feed-back approach to error control in finite element methods: Basic analysis and examples. East-West J Numer Math 1996;4:237–64.

[93] Becker R, Rannacher R. An optimal control approach to a posteriori error estimation in finite element methods. Acta Numer 2001;10:1–102.

[94] Becker R, Heuveline V, Rannacher R. An optimal control approach to adaptivity in computational fluid dynamics. Internat J Numer Methods Fluids 2002;40:105–20.

[95] Heuveline V, Rannacher R. Duality-based adaptivity in the hp-finite element method. J Numer Math 2003;11:95–103.

[96] Venditti DA, Darmofal DL. Adjoint error estimation and grid adaptation for functional outputs: Application to quasi-one-dimensional flow. J Comput Phys 2000;164:204–27.

[97] Venditti DA, Darmofal DL. Grid adaptation for functional outputs: Application to two-dimensional inviscid flows. J Comput Phys 2002;176:40–69.

[98] Venditti DA, Darmofal DL. Anisotropic grid adaptation for functional outputs: Application to two-dimensional viscous flows. J Comput Phys 2003;187:22–46.

[99] Nemec M, Aftosmis MJ. Adjoint error estimation and adaptive refinement for embedded-boundary Cartesian meshes. Paper 2007–4187, AIAA; 2007.

[100] Nemec M, Aftosmis MJ. Adjoint sensitivity computations for an embedded-boundary Cartesian mesh method. J Comput Phys 2008;227(4):2724–42.

[101] Ceze MA, Fidkowski KJ. An anisotropic hp-adaptation framework for functional prediction. AIAA J 2013;51(2):492–509.

[102] Marta AC, Alonso JJ. Discrete adjoint formulation for the ideal MHD equations. Paper 2006-3345, AIAA; 2006.

[103] Marta AC. Rapid development of discrete adjoint solvers with applications to magnetohydrodynamic flow control [Ph.D. thesis], Stanford University; 2007.

[104] Marta AC, Alonso JJ. Toward optimally seeded airflow on hypersonic vehicles using control theory. Comput & Fluids 2010;39(9):1562–74. http://dx.doi.org/10.1016/j.compfluid.2010.05.009.

[105] Narechania N. Output-based error estimation and anisotropic AMR for numerical prediction of steady ideal magnetohydrodynamic flows [Ph.D. thesis], University of Toronto; 2020.

[106] Navon IM, Zou X, Derber J, Sela J. Variational data assimilation with an adiabatic version of the NMC spectral model. Mon Weather Rev 1992;120(7):1433–46.

[107] Brackbill JU, Barnes DC. The effect of nonzero $\nabla \cdot \mathbf{b}$ on the numerical solution of the magnetohydrodynamic equations. J Comput Phys 1980;35:426–30.

[108] Dedner A, Kemm F, Kröner D, Munz C-D, Schnitzer T, Wesenberg M. Hyperbolic divergence cleaning for the MHD equations. J Comput Phys 2002;175:645–73.

[109] Evans CR, Hawley JF. Simulation of magnetohydrodynamic flows: A constrained transport method. Astrophys J 1988;332:659–77.

[110] Scheuerer M, Schlather M. Covariance models for divergence-free and curl-free random vector fields. Stoch Models 2012;28(3):433–51. http://dx.doi.org/10.1080/15326349.2012.699756.

[111] Tóth G. The $\nabla \cdot$ B=0 constraint in shock-capturing magnetohydrodynamics codes. J Comput Phys 2000;161:605–52. http://dx.doi.org/10.1006/jcph.2000.6519.

[112] Brio M, Wu CC. An upwind differencing scheme for the equations of ideal magnetohydrodynamics. J Comput Phys 1988;75:400–22.

[113] Shu C-W, Osher S. Efficient implementation of essentially non-oscillatory shock-capturing schemes: II. J Comput Phys 1989;83:32–78.

[114] Susanto A. High-order finite-volume schemes for magnetohydrodynamics [Ph.D. thesis], University of Waterloo; 2014.