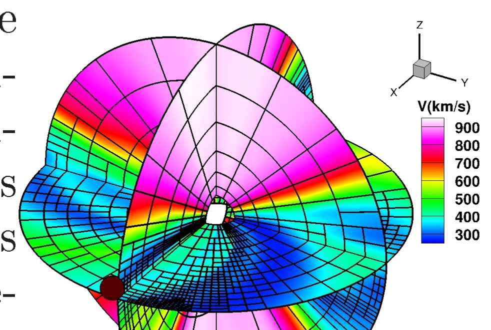
# Toward Data Assimilation of the Solar Wind: Comparison of Variational and Sequential Assimilation for 1D Magnetohydrodynamics Flows

Jose H. Arnal, Clinton P. T. Groth

University of Toronto Institute for Aerospace Studies

jose.arnal@mail.utoronto.ca, groth@utias.utoronto.ca

## Introduction




Due to the potential risks that space weather (SW) events pose on modern technology and infrastructure, there has been increasing interest in physics-based forecasts of the solar wind. Over the last few decades significant effort has been devoted to the development of efficient numerical schemes for space plasmas simulations. More recently, there has been increasing interest in incorporating observational data within SW simulations via data assimilation (DA). In this study, the results of the assimilation of synthetic plasma observations in a one-dimensional ideal MHD initial value problem are presented. Both, variational and sequential DA methods are employed and compared. We thus presents a key milestone in moving towards DA of the three-dimensional solar wind.

## Magnetohydrodynamics

Magnetohydrodynamics is a general mathematical description of electrically conducting fluids and their dynamics. Space plasmas are typically modeled via the **ideal** Magnetohydrodynamics equations (MHD) which assume a perfectly electrically conducting, fully ionized, quasi-neutral, inviscid, ideal gas. In 1D, the ideal MHD equations may be written in the following conservative form

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

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho\mathbf{u} \\ e \\ \mathbf{B} \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho\mathbf{u} \\ \rho\mathbf{u}\mathbf{u} - \mathbf{B}\mathbf{B} + p_T\mathbf{I} \\ (e+p_T)\mathbf{u} - (\mathbf{u}\cdot\mathbf{B})\mathbf{B} \\ \mathbf{B}\mathbf{u} - \mathbf{u}\mathbf{B} \end{bmatrix},$$

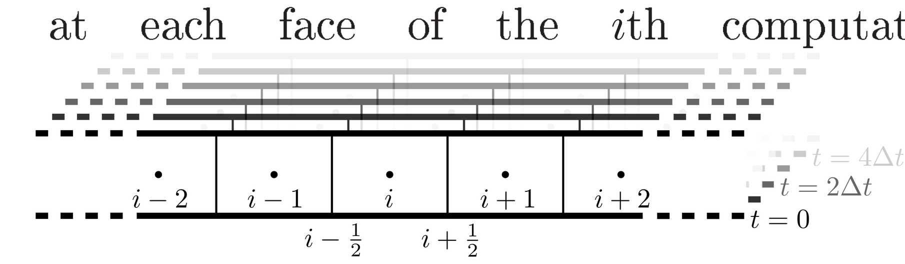
noting that in 1D, $\nabla\cdot\mathbf{B} = 0$ reduces to $B_x =$ constant.

## Finite Volume Method

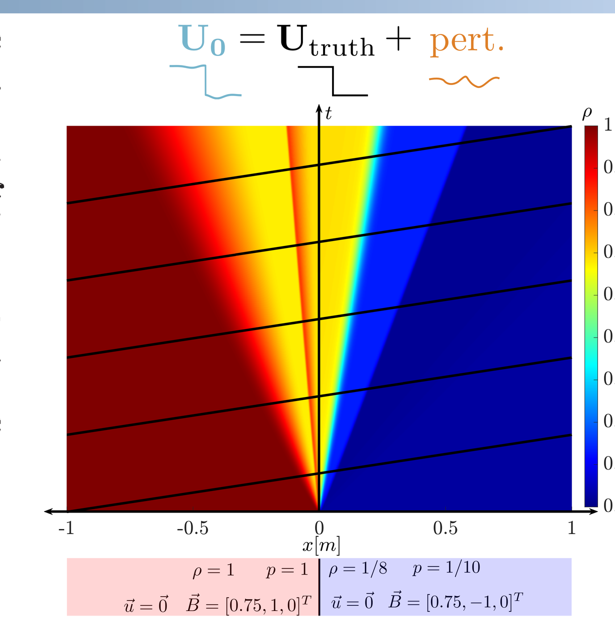
The ideal MHD equations form a coupled system of hyperbolic PDE's. Thus, a standard Godunov-type first-order upwind finite-volume method with the Riemann-solver-based flux function of Powell is used here to obtain numerical solutions on a uniform 1D mesh. The fully discrete form of this scheme is given by

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n + \frac{\Delta t}{\Delta x}[\mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}}],$$

where $\mathbf{F}_{i+\frac{1}{2}}$ is the numerical flux function defined by the approximate solution to the Riemann problem computed at each face of the $i$th computational cell.



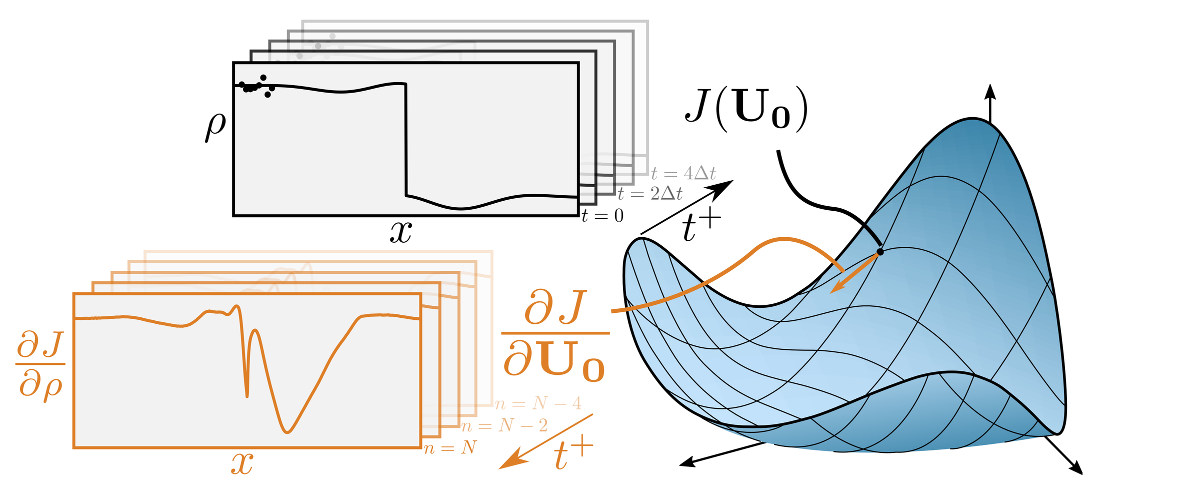

## 1D Brio and Wu Shock Tube Problem




The Brio and Wu shock tube problem is a canonical test for ideal MHD numerical solution schemes. The test comprises of an initial condition where a high-pressure, high-density plasma is separated from a low-density plasma by a thin membrane. The membrane is removed at $t = 0$, resulting in the propagation of various waves.

In this study we consider the Brio-Wu initial condition as the true state. The initial data is generated by perturbing the true state with the addition of random curves. Noisy measurements are generated from $n$ passes of a fictitious observer.

## 4D Variational Data Assimilation




Variational DA poses the assimilation of observations as a minimization problem. A cost function, $J$, is defined which measures the discrepancies between measurements and the model state, as well as differences between the corrected state and the background state. $J$ is minimized by a gradient-based minimization scheme.

$$J = \frac{1}{2}(\mathbf{U}_0 - \mathbf{U}_0^*)^T \boldsymbol{\Sigma}_{\mathbf{U}}^{-1}(\mathbf{U}_0 - \mathbf{U}_0^*) + \frac{1}{2}\sum_{n=0}^{N}(\mathcal{H}(\mathbf{U}^n) - \mathbf{z}^n)^T \boldsymbol{\Sigma}_{\mathbf{z}}^{-1}(\mathcal{H}(\mathbf{U}^n) - \mathbf{z}^n)$$

subject to $\quad \mathbf{R}_i^{*n} = \mathbf{U}_i^{n+1} - \mathbf{U}_i^n + \frac{\Delta t}{\Delta x}[\mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}}] = 0$

This constrained minimization problem can be cast as an unconstrained minimization problem by the method of Lagrange multipliers. We thus define the Lagrangian

$$\mathcal{L} = J + \sum_{n=0}^{N-1}(\boldsymbol{\psi}^{n+1})^T\mathbf{R}^{*n}$$

Taking derivatives w.r.t $\mathbf{U}^n$ and setting them to zero yields the adjoint equations.

$$\boldsymbol{\psi}^N = -\mathbf{H}_n^T\Sigma_{\mathbf{z}}^{-1}(\mathcal{H}(\mathbf{U}^N) - \mathbf{z}^N)$$
$$\boldsymbol{\psi}^n = \mathbf{M}_n^T\boldsymbol{\psi}^{n+1} - \mathbf{H}_n^T\Sigma_{\mathbf{z}}^{-1}(\mathcal{H}(\mathbf{U}^n) - \mathbf{z}^n) \quad \forall n \in [N-1, \cdots, 1]$$

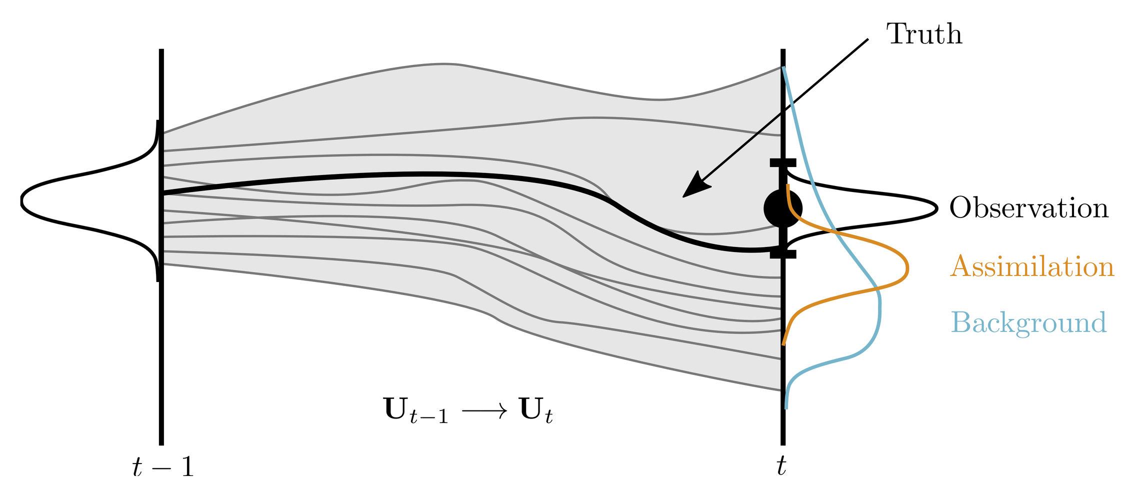
Where

$$\mathcal{M}_i^n = \mathbf{U}_i^n + \frac{\Delta t}{\Delta x}[\mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}}], \ \ \mathbf{M}^n = \frac{\partial\mathcal{M}(\mathbf{U}^n)}{\partial\mathbf{U}^n}, \ \ \mathbf{H}^n = \frac{\partial\mathcal{H}(\mathbf{U}^n)}{\partial\mathbf{U}^n}$$

It can be shown that the gradient of $J$ w.r.t $\mathbf{U_0}$ is equal to:

$$\frac{\partial J}{\partial\mathbf{U}_0} = \boldsymbol{\Sigma}_{\mathbf{U}}^{-1}(\mathbf{U}_0 - \mathbf{U}_0^*) - \boldsymbol{\psi}^0$$

## Ensemble Kalman Filter




Sequential DA algorithms follow a Bayesian framework whereby the system state is updated successively as observations become available in time. Once the state has been updated by the assimilation of observations, it is integrated forward in time until the next set of observations. The Ensemble Kalman Filter (EnKF), is a sequential algorithm that extends the Kalman Filter (KF) to non-linear models. The KF algorithm defines the Kalman gain, $\mathbf{K}^n$, which minimizes the variance of the corrected state, $\hat{\mathbf{U}}^n$, given Gaussian error statistics.

$$\mathbf{K}^n = \boldsymbol{\Sigma}_{\mathbf{U}}^n\mathbf{H}^T(\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{U}}^n\mathbf{H} + \boldsymbol{\Sigma}_{\mathbf{z}})^{-1}$$
$$\hat{\mathbf{U}}^n = \overline{\mathbf{U}^n} + \mathbf{K}^n(\mathbf{z}^n - \mathcal{H}(\overline{\mathbf{U}^n}))$$

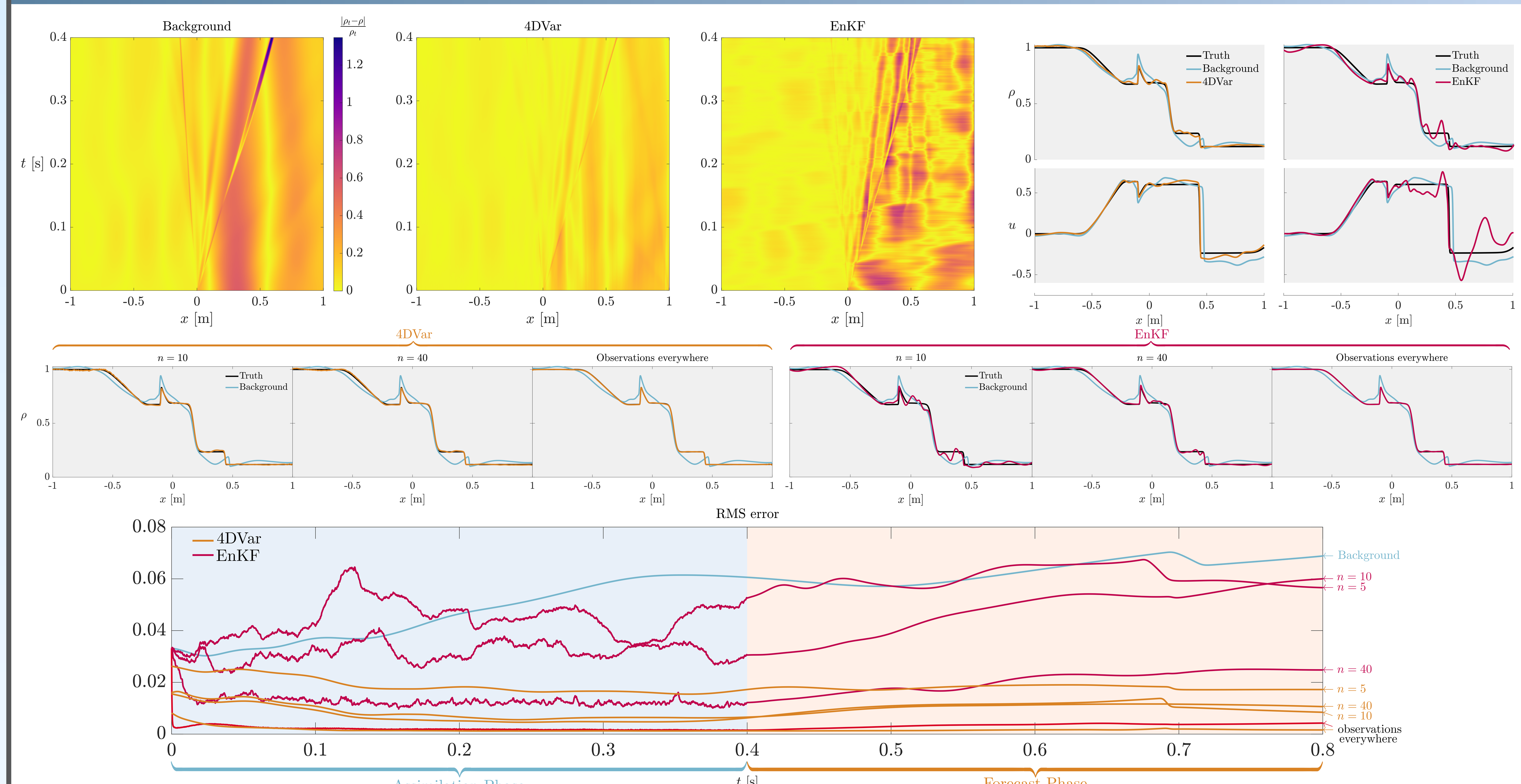
The EnKF estimates the model error covariance, $\boldsymbol{\Sigma}_{\mathbf{U}}^n$, via a monte carlo integration requiring $S$ model runs. To generate the ensemble, the state is perturbed by the addition of random curves.

$$\overline{\mathbf{U}^n} = \frac{1}{S}\sum_{s=1}^{S}\mathbf{U}^{n,s}$$

$$\boldsymbol{\Sigma}_{\mathbf{U}}^n = \frac{1}{S-1}\sum_{s=1}^{S}(\mathbf{U}^{n,s} - \overline{\mathbf{U}^n})(\mathbf{U}^{n,s} - \overline{\mathbf{U}^n})^T$$

This correction takes place whenever an observation is available. If observations are not available at the present time step, the ensemble of model states are integrated forward in time.

## Comparison of 4DVar and EnKF with Synthetic Shock Tube Data




## References

1. K.J.H. Law, A.M. Stuart, K.C. Zygalakis. Data Assimilation: A Mathematical Introduction. 2015
2. G. Evensen. The ensemble Kalman filter: Theoretical formulation and practical implementation. Ocean dynamics, 2003
3. Narechania *et al.* An integrated data-driven solar wind – CME numerical framework for space weather forecasting. Journal of Space Weather and Space Climate, 2021
4. M. Brio and C. C. Wu. An upwind differencing scheme for the equations of ideal magnetohydrodynamics. Journal of Computational Physics, 1988