

Data Assimilation for Magnetohydrodynamics with a Zero-Divergence Constraint on the Magnetic Field

Bruno O. S. Teixeira^{*†}, Aaron Ridley[‡], Leonardo A. B. Tôrres^{*}, Luis A. Aguirre^{*}, and Dennis S. Bernstein[†]

Abstract—We address the problem of enforcing the zero-divergence constraint in data assimilation for two-dimensional magnetohydrodynamic flow. Using a finite-volume computational model, we compare the performance of the localized unscented Kalman filter with the localized equality-constrained unscented and projected localized unscented Kalman filters, which enforce the zero-divergence constraint.

I. INTRODUCTION

Magnetohydrodynamics (MHD) describes plasma dynamics as a fluid moving under the influence of electromagnetic and pressure-gradient forces [6]. MHD involves continuous-time coupled partial differential equations for which it is generally difficult to obtain closed-form solutions. Computational methods for MHD are based on finite-volume spatial and temporal discretization schemes [10]. In addition to difficulties associated with high nonlinearity and large dimension of the discretized equations, there is the concern about maintaining the zero-divergence property of the magnetic field [2, 13] as determined by the finite-difference approximation of the corresponding continuum quantity.

Three approaches are popular for handling the zero-divergence constraint in MHD simulation, namely, (i) the eight-wave formulation [13], (ii) constrained transport and central difference discretizations [13], and (iii) the projection scheme [2]. The first approach employs a nonconservative formulation of the MHD equations, where terms proportional to the divergence are added, resulting in a more numerically robust discretized model. Constrained transport and central difference methods use special finite-difference discretizations to maintain the zero-divergence property. In the projection approach, the numerical solution of the nonzero-divergence finite-volume discretization scheme is projected onto the subspace of zero divergence.

Data assimilation for MHD is of interest for space weather forecasting applications [3, 7]. Similar to the simulation problem, data assimilation for MHD [3, 4] may violate the zero-divergence constraint, even if the injected data and the model satisfy this equality constraint. However, zero-divergence data-assimilation algorithms have not been applied to MHD.

This research was supported by the National Council for Scientific and Technological Development (CNPq), Brazil, and by the National Science Foundation Information Technology Research Initiative, through Grants ATM-0325332 and CNS-0539053 to the University of Michigan, USA.

^{*}Department of Electronic Engineering, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil, {brunoot, torres, aguirre}@cpdee.ufmg.br.

[†]Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI, USA, dsbaero@umich.edu.

[‡]Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, MI, USA, ridley@umich.edu.

The contribution of the present paper is to demonstrate two data-assimilation techniques that enforce the zero-divergence constraint in the magnetic field.

To begin, we briefly review the field-interpolated central difference [13] and projection [2, 13] approaches for MHD simulation. Next, for data assimilation, to enforce the zero-divergence constraint while estimating the flow variables, we present the localized equality-constrained unscented Kalman filter (LECUKF), a reduced-order application of the equality-constrained Kalman filter [11, 12], as well as the projected localized unscented Kalman filter (PLUKF), a reduced-order extension of the projected unscented Kalman filter [11, 12]. Results are compared with the (unconstrained) localized unscented Kalman filter (LUKF) [3, 4], whose estimates do not satisfy the zero-divergence constraint.

We use a two-dimensional numerical example to illustrate and compare LUKF, LECUKF, and PLUKF. First, we briefly compare MHD simulated data regarding the zero-divergence property using the unconstrained second-order Russanov scheme [10] and the field-interpolated central difference and projection approaches. Then we use zero-divergence magnetic field data and the field-interpolated central difference model to perform data assimilation using LUKF and show that the estimates do not satisfy the zero-divergence property. We employ LECUKF to guarantee that the estimates in the localized region have zero divergence within numerical precision. Finally, we employ PLUKF to obtain estimates in the full grid satisfying zero divergence, but without feeding these constrained estimates back into the next time step.

II. MHD SIMULATION

The ideal MHD equations are given by [6]

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \vec{v}), \quad (2.1)$$

$$\frac{\partial \rho \vec{v}}{\partial t} = -\nabla \cdot \left[\rho \vec{v} \vec{v}^T + \left(p + \frac{1}{2} |\vec{B}|^2 \right) I_{3 \times 3} - \vec{B} \vec{B}^T \right], \quad (2.2)$$

$$\frac{\partial \vec{B}}{\partial t} = -\nabla \cdot \left(\rho \vec{v} \vec{B}^T - \vec{B} \vec{v}^T \right), \quad (2.3)$$

$$\frac{\partial \mathcal{E}}{\partial t} = -\nabla \cdot \left[\left(\mathcal{E} + p + \frac{1}{2} |\vec{B}|^2 \right) \vec{v} - \vec{B} (\vec{v} \cdot \vec{B}) \right], \quad (2.4)$$

together with the zero-divergence constraint

$$\nabla \cdot \vec{B} = 0, \quad (2.5)$$

where $\rho > 0$ is the mass density, $\vec{v} \triangleq [v_x \ v_y \ v_z]^T \in \mathbb{R}^3$ is the velocity, $\vec{m} \triangleq \rho \vec{v} \in \mathbb{R}^3$ is the momentum, $\vec{B} \triangleq [B_x \ B_y \ B_z]^T \in \mathbb{R}^3$ is the magnetic field, $\mathcal{E} > 0$ is the

total energy, and the hydrodynamic pressure $p \in \mathbb{R}$ is given by $p = (\gamma - 1) \left(\mathcal{E} - \frac{1}{2} \rho |\vec{v}|^2 - \frac{1}{2} |\vec{B}|^2 \right)$, where γ is the specific heat ratio. For simplicity we consider only the two-dimensional case, where $v_z = 0$ and $B_z = 0$. However, the techniques that we employ can readily be extended to the three-dimensional case.

We apply finite-volume-based spatial and temporal discretization to (2.1)-(2.4) [10]. Assume that the channel consists of $n_x \times n_y$ identical cells. For all $i = 1, \dots, n_x$, $j = 1, \dots, n_y$, let $\rho^{[i,j]}$, $m_x^{[i,j]}$, $m_y^{[i,j]}$, $B_x^{[i,j]}$, $B_y^{[i,j]}$, and $\mathcal{E}^{[i,j]}$ denote, respectively, the density, the momentum components in the x and y directions, the magnetic field components in the x and y directions, and the energy in cell $[i, j]$. Define $U^{[i,j]} \in \mathbb{R}^6$ by

$$U^{[i,j]} = \left[\rho^{[i,j]} \quad m_x^{[i,j]} \quad m_y^{[i,j]} \quad B_x^{[i,j]} \quad B_y^{[i,j]} \quad \mathcal{E}^{[i,j]} \right]^T. \quad (2.6)$$

We use a second-order Rusanov scheme [10] to discretize (2.1)-(2.4) and obtain a discrete-time model that enables us to update the flow variables at the center of each cell. The discrete-time state update equation [10] is given by

$$U_k^{[i,j]} = U_{k-1}^{[i,j]} - \frac{T}{\Delta x} \left[\bar{F}_{\text{Rus},k-1}^{[i,j]} - \bar{F}_{\text{Rus},k-1}^{[i-1,j]} \right] - \frac{T}{\Delta y} \left[\bar{F}_{\text{Rus},k-1}^{[i,j]} - \bar{F}_{\text{Rus},k-1}^{[i,j-1]} \right], \quad (2.7)$$

where $T > 0$ is the sampling time, Δx and Δy are the width and height of each cell, respectively, and $\bar{F}_{\text{Rus},k-1}^{[i,j]}$ is a nonlinear function of $U_{k-1}^{[i-1,j-1]}, \dots, U_{k-1}^{[i+2,j+2]}$; see [10]. Hence, $U_k^{[i,j]}$ depends on $U_{k-1}^{[i-2,j-2]}, \dots, U_{k-1}^{[i+2,j+2]}$, as expected for a spatially second-order scheme. Henceforth, we refer to (2.7) as the *base* scheme for MHD simulation. Likewise, for $i = 2, \dots, n_x - 1$, $j = 2, \dots, n_y - 1$, define

$$\nabla B_k^{[i,j]} \triangleq \frac{1}{\Delta x} \nabla B_{x,k}^{[i,j]} + \frac{1}{\Delta y} \nabla B_{y,k}^{[i,j]}, \quad (2.8)$$

where $\nabla B_{x,k}^{[i,j]} \triangleq \frac{B_{x,k}^{[i+1,j]} - B_{x,k}^{[i-1,j]}}{2}$ and $\nabla B_{y,k}^{[i,j]} \triangleq \frac{B_{y,k}^{[i,j+1]} - B_{y,k}^{[i,j-1]}}{2}$, as an approximation for $\nabla \cdot \vec{B}$.

Next, define the state vector $x_k \in \mathbb{R}^n$, where $n \triangleq 6(n_x - 4)(n_y - 4)$, by

$$x_k \triangleq \left[(U_k^{[3,3]})^T \dots (U_k^{[3,n_y-2]})^T \dots (U_k^{[n_x-2,3]})^T \dots (U_k^{[n_x-2,n_y-2]})^T \right]^T. \quad (2.9)$$

Let the input vector $u_{k-1} \in \mathbb{R}^p$, where $p \triangleq 24(n_x + n_y - 4)$, denote the boundary conditions for the left, right, bottom, and right cells, that is,

$$u_{k-1} \triangleq \left[u_{1,k-1}^T \quad u_{2,k-1}^T \right]^T, \quad (2.10)$$

where $p_1 \triangleq 12(n_x + n_y - 2)$, $p_2 \triangleq p - p_1$, and $u_{1,k-1} \in \mathbb{R}^{p_1}$, $u_{2,k-1} \in \mathbb{R}^{p_2}$ are given by

$$u_{1,k-1} \triangleq \left[(U_{k-1}^{[1,1;n_y]})^T (U_{k-1}^{[n_x,1;n_y]})^T (U_{k-1}^{[2;n_x-1,1]})^T (U_{k-1}^{[2;n_x-1,n_y]})^T \right]^T, \quad (2.11)$$

$$u_{2,k-1} \triangleq \left[(U_{k-1}^{[2,2;n_y-1]})^T (U_{k-1}^{[n_x-1,2;n_y-1]})^T \right]^T$$

$$(U_{k-1}^{[3;n_x-2,2]})^T (U_{k-1}^{[3;n_x-2,n_y-1]})^T \Big]. \quad (2.12)$$

We now rewrite (2.7) as the nonlinear discrete-time model

$$x_k = f(x_{k-1}, u_{k-1}), \quad (2.13)$$

where $f : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^n$. We assume zero-divergence boundary conditions, that is, for $i = 1, 2, n_x - 1, n_x$ and $j = 1, 2, n_y - 1, n_y$, we assume that

$$\nabla B_k^{[i,j]} = 0 \quad (2.14)$$

is satisfied. However, the discretized model (2.13) does not guarantee that (2.14) is satisfied for all $k \geq 1$, $i = 3, \dots, n_x - 2$, and $j = 3, \dots, n_y - 2$.

To account for unknown disturbances $w_{k-1} \in \mathbb{R}^q$, let (2.13) be rewritten as the *truth model*

$$x_k = f(x_{k-1}, u_{k-1}) + G_{k-1} w_{k-1}, \quad (2.15)$$

where $G_{k-1} \in \mathbb{R}^{n \times q}$ and w_{k-1} is a zero-mean, white, Gaussian process noise with covariance Q_{k-1} . Like (2.13), $G_{k-1} w_{k-1}$ in (2.15) can violate (2.14).

III. ZERO-DIVERGENCE MHD SIMULATION

A. Field-Interpolated Central Difference Scheme

Let $\Omega_{z,k}^{[i,j]} \in \mathbb{R}$ denote the time-centered approximation of the z -component of the electric field at cell $[i, j]$,

$$\Omega_{z,k}^{[i,j]} = \text{row}_3 \left(-\frac{\vec{v}_{k-1}^{[i,j]} \times \vec{B}_{k-1}^{[i,j],\text{cd}} + \vec{v}_k^{[i,j]} \times \vec{B}_k^{[i,j]}}{2} \right), \quad (3.1)$$

where $\vec{v}_k^{[i,j]} \triangleq \begin{bmatrix} v_{x,k}^{[i,j]} & v_{y,k}^{[i,j]} & 0 \end{bmatrix}^T$ and $\vec{B}_k^{[i,j]} \triangleq \begin{bmatrix} B_{x,k}^{[i,j]} & B_{y,k}^{[i,j]} & 0 \end{bmatrix}^T$ are given by the base scheme (2.7), and, for $i = 3, \dots, n_x - 2$ and $j = 3, \dots, n_y - 2$, the components of $\vec{B}_{k-1}^{[i,j],\text{cd}}$ are recursively updated as

$$B_{x,k}^{[i,j],\text{cd}} = B_{x,k-1}^{[i,j],\text{cd}} - \frac{T}{2\Delta y} \left(\Omega_{z,k}^{[i,j+1]} - \Omega_{z,k}^{[i,j-1]} \right), \quad (3.2)$$

$$B_{y,k}^{[i,j],\text{cd}} = B_{y,k-1}^{[i,j],\text{cd}} + \frac{T}{2\Delta x} \left(\Omega_{z,k}^{[i+1,j]} - \Omega_{z,k}^{[i-1,j]} \right). \quad (3.3)$$

Then, by appending (3.1)-(3.3) to (2.7), we have the field-interpolated *central-difference* scheme (CD) [13], which, in order to satisfy (2.14), employs central differencing for the induction equation (2.4) on the original grid.

B. Projection Scheme

Let $B_k^{[i,j]} \triangleq \begin{bmatrix} B_{x,k}^{[i,j]} & B_{y,k}^{[i,j]} \end{bmatrix}$ and $B_k \triangleq \begin{bmatrix} B_k^{[2,2]} \dots B_k^{[2,n_x-1]} \dots B_k^{[n_x-1,1]} \dots B_k^{[n_x-1,n_y-1]} \end{bmatrix}^T \in \mathbb{R}^{n_B}$, where $n_B \triangleq (n + p_2)/6$, denotes the numerical values of the magnetic field provided by (2.7) plus the cells of the boundary conditions given by (2.12) at time k . Furthermore, for $E \in \mathbb{R}^{(n/6) \times n_B}$, rewrite (2.8) in the matrix form

$$\nabla B_k \triangleq EB_k. \quad (3.4)$$

The *projection* scheme [2, 13] projects B_k orthogonally onto the subspace of zero-divergence solutions by minimizing the cost function $J(B_k^{\text{P}}) \triangleq (B_k^{\text{P}} - B_k)^T (B_k^{\text{P}} - B_k)$ subject to $EB_k^{\text{P}} = 0_{(n/6) \times 1}$. The solution is given by

$$B_k^{\text{P}} = B_k + K(0_{(n/6) \times 1} - EB_k) = \mathcal{P}B_k, \quad (3.5)$$

where the gain matrix is given by $K \triangleq E^T(EE^T)^{-1} \in \mathbb{R}^{n_B \times (n/6)}$ and the orthogonal projector is given by $\mathcal{P} \triangleq I_{n_B \times n_B} - KE \in \mathbb{R}^{n_B \times n_B}$. The projected magnetic field B_k^P is used at $k+1$. Note that K and \mathcal{P} need to be computed only once.

This direct approach (3.5) is restricted to moderate grid sizes due to large memory demand [13]. Also, density, moment components, and energy are not affected by the projection scheme at the current time step.

IV. MHD DATA ASSIMILATION

For the process model (2.15), we assume that, for all $k \geq 1$, in addition to inputs $u_{k-1} \in \mathbb{R}^p$ given by (2.10), measurements $y_k \in \mathbb{R}^m$ of flow variables are available in certain cells. We represent y_k as

$$y_k = C_k x_k + v_k, \quad (4.1)$$

where $C_k \in \mathbb{R}^{m \times n}$ and $v_k \in \mathbb{R}^m$ is a zero-mean, white, Gaussian measurement noise with covariance R_k . Assume that w_{k-1} and v_k are mutually independent. Also, rewrite the zero-divergence constraint (2.14) as

$$D \begin{bmatrix} x_k \\ u_{2,k} \end{bmatrix} = 0_{(n/6) \times 1}, \quad (4.2)$$

where $D \in \mathbb{R}^{(n/6) \times (n+p_2)}$ and $u_{2,k}$ is given by (2.12). Assume that we know only the initial estimate $\hat{x}_{0|0}$ and the error-covariance $P_{0|0}^{xx} \triangleq \mathcal{E} \left[(x_0 - \hat{x}_{0|0})(x_0 - \hat{x}_{0|0})^T \right]$ of the initial state vector x_0 , which is assumed to be Gaussian. Next, define the profit function $J(x_k) \triangleq \rho(x_k | (y_1, \dots, y_k))$, which is the conditional probability density function of x_k given the past and present measured data y_1, \dots, y_k . Under the stated assumptions, the maximization of $J(x_k)$ subject to (4.2) is the *equality-constrained data-assimilation* problem.

The solution to this problem is complicated by the existence of the zero-divergence constraint and by the fact that $\rho(x_k | (y_1, \dots, y_k))$ is not completely characterized by its first-order and second-order moments since (2.15) is nonlinear [5]. Therefore, approximate unconstrained solutions based on the Kalman filter (KF) [8] are commonly used, for example, the unscented Kalman filter (UKF) [9]. Moreover, since the dimension of x_k for MHD applications is large, reduced-order algorithms are generally employed [4].

To obtain a reduced-order estimator, we partition $x_k \in \mathbb{R}^n$ as $x_k = \begin{bmatrix} x_{L,k}^T & x_{E,k}^T \end{bmatrix}^T$, where $x_{L,k} \in \mathbb{R}^{n_L}$ accounts for the localized region where measurements are available, while $x_{E,k} \in \mathbb{R}^{n_E}$ accounts for the exterior part of the grid without measurements. In this case, (4.1) can be expressed as

$$y_k = C_{L,k} x_{L,k} + v_k, \quad (4.3)$$

where $C_{L,k} \in \mathbb{R}^{m \times n_L}$ is formed by the columns of C_k associated with $x_{L,k}$. Similarly, from (4.2), we have

$$D_L x_{L,k} = 0_{(n_L/6) \times 1}, \quad (4.4)$$

where $D_L \in \mathbb{R}^{(n_L/6) \times n_L}$. Let $n_{x_{L0}}$, $n_{x_{Lf}}$, $n_{y_{L0}}$, and $n_{y_{Lf}}$ denote the coordinates of the localized region of the grid. The objective is to directly inject the measurement data y_k

into only the states corresponding to the estimate of $x_{L,k}$. We thus use the localized (unconstrained) unscented Kalman filter (LUKF) [4] to provide a suboptimal solution to the data-assimilation problem. Moreover, to enforce the zero-divergence constraint into the estimates of the localized region, we present the localized equality-constrained unscented Kalman filter (LECUKF). Also, we present the projected localized unscented Kalman filter (PLUKF) to enforce the zero-divergence constraint in the full grid.

V. LOCALIZED UNSCENTED KALMAN FILTER

Instead of analytically or numerically linearizing (2.15) and using the KF equations [8], UKF employs the unscented transform (UT) [9], which is a numerical procedure for approximating the mean and covariance of a random vector obtained from a nonlinear transformation.

Regarding system given by (2.15) and (4.3), LUKF is a two-step algorithm, whose *forecast* equations are given by

$$\begin{aligned} \mathcal{X}_{L,k-1|k-1} &\triangleq \hat{x}_{L,k-1|k-1} \mathbf{1}_{1 \times (2n_L+1)} + \sqrt{(n_L + \lambda)} \\ &\times \left[0_{n_L \times 1} \left(P_{L,k-1|k-1}^{xx} \right)^{1/2} - \left(P_{L,k-1|k-1}^{xx} \right)^{1/2} \right], \end{aligned} \quad (5.1)$$

$$\mathcal{X}_{k-1|k-1} \triangleq \begin{bmatrix} \mathcal{X}_{L,k-1|k-1} \\ \hat{x}_{E,k-1|k-1} \mathbf{1}_{1 \times (2n_L+1)} \end{bmatrix}, \quad (5.2)$$

$$\text{col}_i(\mathcal{X}_{k|k-1}) \triangleq f(\text{col}_i(\mathcal{X}_{k-1|k-1}), u_{k-1}), \quad i = 0, \dots, 2n_L, \quad (5.3)$$

$$\hat{x}_{k|k-1} \triangleq \sum_{i=0}^{2n_L} \gamma_i^{(m)} \text{col}_i(\mathcal{X}_{k|k-1}), \quad (5.4)$$

$$P_{L,k|k-1}^{xx} \triangleq \sum_{i=0}^{2n_L} \gamma_i^{(c)} [\text{col}_i(\mathcal{X}_{L,k|k-1}) - \hat{x}_{L,k|k-1}] \quad (5.5)$$

$$\times [\text{col}_i(\mathcal{X}_{L,k|k-1}) - \hat{x}_{L,k|k-1}]^T + G_{L,k-1} Q_{L,k-1} G_{L,k-1}^T,$$

$$\hat{y}_{k|k-1} \triangleq C_{L,k} \hat{x}_{L,k|k-1}, \quad (5.6)$$

$$P_{L,k|k-1}^{yy} \triangleq C_{L,k} P_{L,k|k-1}^{xx} C_{L,k}^T + R_k, \quad (5.7)$$

$$P_{L,k|k-1}^{xy} \triangleq P_{L,k|k-1}^{xx} C_{L,k}^T, \quad (5.8)$$

where, for $i = 1, \dots, n_L$, the weights are given by

$$\begin{cases} \gamma_0^{(m)} \triangleq \frac{\lambda}{n_L + \lambda}, \\ \gamma_0^{(c)} \triangleq \frac{\lambda}{n_L + \lambda} + 1 - \alpha^2 + \beta, \\ \gamma_i^{(m)} \triangleq \gamma_i^{(c)} \triangleq \gamma_{i+n_L}^{(m)} \triangleq \gamma_{i+n_L}^{(c)} \triangleq \frac{1}{2(n_L + \lambda)}, \end{cases} \quad (5.9)$$

$(\cdot)^{1/2}$ is the Cholesky square root, $0 < \alpha \leq 1$, $\beta \geq 0$, $\kappa \geq 0$, and $\lambda \triangleq \alpha^2(\kappa + n_L) - n_L > -n_L$, $\begin{bmatrix} \mathcal{X}_{L,k|k-1} \\ \mathcal{X}_{E,k|k-1} \end{bmatrix} \triangleq$

$\mathcal{X}_{k|k-1}$, $\begin{bmatrix} \hat{x}_{L,k|k-1} \\ \hat{x}_{E,k|k-1} \end{bmatrix} \triangleq \hat{x}_{k|k-1}$, $P_{L,k-1|k-1}^{xx}$ is the localized data-assimilation error covariance, $P_{L,k|k-1}^{xx}$ is the localized forecast error covariance, $P_{L,k|k-1}^{yy}$ is the localized innovation covariance, and $P_{L,k|k-1}^{xy}$ is the localized cross covariance. We set $\alpha = 0.6$, $\beta = 0$, and $\kappa = 0$ [3]. The notation $\hat{x}_{k|k-1}$ indicates an estimate of x_k at time k based on information available up to and including time $k-1$. Likewise, \hat{x}_k indicates an estimate of x_k at time k using information available up to and including time k .

The *data-assimilation* equations are given by

$$K_{L,k} = P_{L,k|k-1}^{xy} (P_{L,k|k-1}^{yy})^{-1}, \quad (5.10)$$

$$\hat{x}_{L,k|k} = \hat{x}_{L,k|k-1} + K_{L,k} (y_k - \hat{y}_{k|k-1}), \quad (5.11)$$

$$\hat{x}_{E,k|k} = \hat{x}_{E,k|k-1}, \quad (5.12)$$

$$P_{L,k|k}^{xx} = P_{L,k|k-1}^{xx} - K_{L,k} P_{L,k|k-1}^{yy} K_{L,k}^T, \quad (5.13)$$

where $K_{L,k} \in \mathbb{R}^{n_L \times m}$ is the localized Kalman gain. Model information is used during the forecast step, while measurement data are injected into the estimates during the data-assimilation step.

VI. LOCALIZED EQUALITY-CONSTRAINED UKF

Now we consider LECUKF to provide a suboptimal solution to the equality-constrained data-assimilation problem. LECUKF is based on the equality-constrained unscented Kalman filter (EUKF) [11, 12], which is a nonlinear extension of the equality-constrained Kalman filter (ECKF) [12]

Regarding the system given by (2.15), (4.3), and (4.4), PLUKF is a three-step algorithm, whose *forecast* equations are given by

$$\mathcal{X}_{L,k-1|k-1}^p \triangleq \hat{x}_{L,k-1|k-1}^p \mathbf{1}_{1 \times (2n_L+1)} + \sqrt{(n_L + \lambda)} \quad (6.1)$$

$$\times \left[0_{n_L \times 1} \left(P_{L,k-1|k-1}^{xx} \right)^{1/2} - \left(P_{L,k-1|k-1}^{xpp} \right)^{1/2} \right],$$

$$\mathcal{X}_{k-1|k-1} = \begin{bmatrix} \mathcal{X}_{L,k-1|k-1}^p \\ \hat{x}_{E,k-1|k-1} \mathbf{1}_{1 \times (2n_L+1)} \end{bmatrix}, \quad (6.2)$$

together with (5.3)-(5.8), where $\hat{x}_{L,k|k}^p \in \mathbb{R}^{n_L}$ is the projected localized state vector and $P_{L,k|k}^{xpp}$ is the localized projected error covariance. The *data-assimilation* equations are given by (5.10)-(5.13). Finally, the zero-divergence constraint is enforced during the *projection* step whose equations are given by

$$K_{L,k}^p = P_{L,k|k}^{xx} D_L^T (D_L P_{L,k|k}^{xx} D_L^T)^{-1}, \quad (6.3)$$

$$\mathcal{P}_{L,k} = (I_{n_L \times n_L} - K_{L,k}^p D_L), \quad (6.4)$$

$$\hat{x}_{L,k|k}^p = \mathcal{P}_{L,k} \hat{x}_{L,k|k}, \quad (6.5)$$

$$P_{L,k|k}^{xpp} = \mathcal{P}_{L,k} P_{L,k|k}^{xx}, \quad (6.6)$$

where $\mathcal{P}_{L,k} \in \mathbb{R}^{n_L \times n_L}$ is an oblique projector whose range is the null space of D_L . For the general case of a nonlinear observation model and a nonlinear equality constraint, the ECUKF equations are presented in [11].

VII. PROJECTED LOCALIZED UKF

PLUKF is based on the projected unscented Kalman filter (PUKF) [11, 12]. Regarding the system given by (2.15), (4.3), and (4.2), PLUKF is a three-step algorithm, whose *forecast* equations are given by (5.1)-(5.8), whose *data-assimilation* equations are given by (5.10)-(5.13), and whose *projection* step is given by

$$K^p = D^T (D D^T)^{-1}, \quad (7.1)$$

$$\mathcal{P} = (I_{n \times n} - K^p D), \quad (7.2)$$

$$\hat{x}_{k|k}^p = \mathcal{P} \hat{x}_{k|k}. \quad (7.3)$$

Note that (7.1)-(7.2) need to be computed only once.

Note that LECUKF enforces the zero-divergence constraint only in the localized region because it requires the calculation of the data-assimilation error covariance; see (6.3). On the other hand, PLUKF performs projection on the full grid. However, unlike LECUKF, PLUKF does not

recursively feed the projected estimate $\hat{x}_{k|k}^p$ back in the forecast step.

VIII. SIMULATION NUMERICAL EXAMPLE

We consider a two-dimensional grid with $n_x = 64$, $n_y = 24$, $\Delta x = 1$, $\Delta y = 1$, and $T = 0.01$ s. We set as initial conditions $U_0^{[i,j]} = [2 \ 10 \ 0 \ 0 \ 1 \ 1]^T$, $i = 1, \dots, n_x$, $j = 1, \dots, n_y$, which represent supersonic flow. For a given cell $[i, j]$, define $G^{[i,j]} \triangleq \text{diag}(0.01, 0.05, 0.05, 0, 0, 0.05)$. We consider the truth model (2.15), where we set $Q_{k-1} = I_{n \times n}$ and

$$G_{k-1} = \text{diag} \left(0_{6 \times 6}, \dots, G^{[14,30]}, \dots, 0_{6 \times 6}, \dots, G^{[14,32]}, \dots, \right. \\ \left. 0_{6 \times 6}, \dots, G^{[14,34]}, \dots, 0_{6 \times 6}, \dots, G^{[15,32]}, \dots, 0_{6 \times 6} \right).$$

Moreover, u_{k-1} (2.10) is defined as follows. For all $k \geq 1$, we assume floating boundary conditions for the bottom and top cells, that is, $U_{k-1}^{[i,j]} = U_{k-1}^{[i,3]}$, $i = 3, \dots, n_x - 2$, $j = 1, 2$, and $U_{k-1}^{[i,j]} = U_{k-1}^{[i, n_y - 2]}$, $i = 3, \dots, n_x - 2$, $j = n_y - 1, n_y$. For the left cells, we set constant boundary conditions $U_{k-1}^{[i,j]} = U_0^{[i,j]}$, $i = 1, 2$, $j = 1, \dots, n_y$. Finally, for the right cells, we set floating boundary conditions $U_{k-1}^{[i,j]} = U_{k-1}^{[n_x - 2, j]}$, $i = n_x - 1, n_x$, $j = 3, \dots, n_y - 2$, except for

$$m_{x,k-1}^{[i,j]} = \begin{cases} -m_{x,k-1}^{[n_x - 2, j]}, & \text{for } i = n_x - 1, n_x, \\ & j = n_y/2 - 1, \dots, n_y/2 + 1, \\ m_{x,k-1}^{[n_x - 2, j]}, & \text{for } i = n_x - 1, n_x, \\ & j = 3, \dots, n_y/2, n_y/2 + 2, \dots, n_y - 2, \end{cases}$$

for which reflective boundary conditions are defined to create a bowshock.

Figure 1 shows the y components of the magnetic field at $kT = 15$ s for three simulated cases, namely, (i) using second-order Rusanov scheme as the base scheme (Base), (ii) combining the base scheme with the central difference scheme (CD), and (iii) combining the base scheme with the projection scheme (Proj). For $i = 3, \dots, n_x - 2$ and $j = 3, \dots, n_y - 2$, data from cases (ii) and (iii) satisfy (2.14). Though a bowshock wave is observed in all cases, the results are different, especially for case (iii). Note that, since the three numerical methods mentioned above provide approximate solutions to the MHD equations, we do not have a truth model for this example to compare our results with.

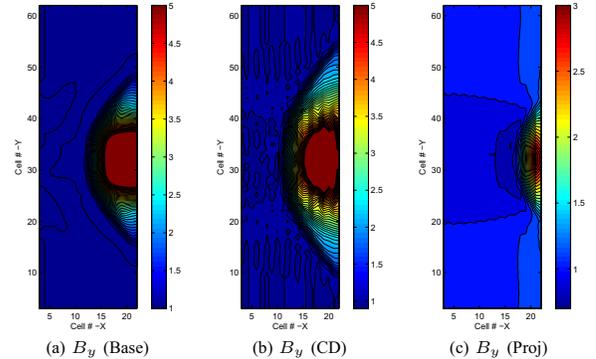


Fig. 1. B_y components of magnetic field at $kT = 15$ s obtained from two-dimensional MHD simulation using the base (Base), central difference (CD), and projection (Proj) approaches.

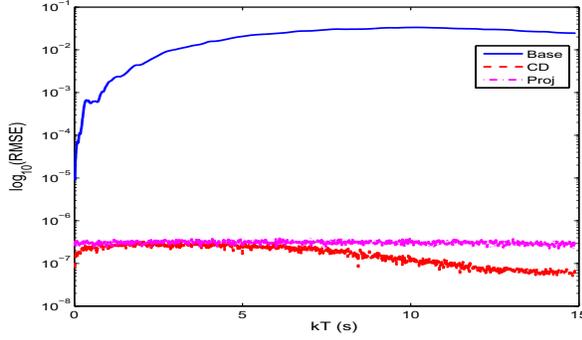


Fig. 2. RMSE index (8.1) evolving with time kT for simulated data using the base (Base), central difference (CD), and projection (Proj) approaches.

The root-mean-square error (RMSE) index is defined as

$$\text{RMSE}_k \triangleq \sqrt{\frac{1}{(i_f - i_0 + 1)(j_f - j_0 + 1)} \sum_{i=i_0}^{i_f} \sum_{j=j_0}^{j_f} \left(\frac{\nabla B_k^{[i,j]}}{\|\bar{B}_k^{[i,j]}\|_2} \right)^2}, \quad (8.1)$$

where the indices i_0 , i_f , j_0 , j_f indicate which part of the grid is used to evaluate RMSE to quantify the mean error associated with the zero-divergence constraint on the grid. Figure 2 shows RMSE_k as a function of time for $i_0 = 3$, $i_f = n_x - 2$, $j_0 = 3$, $j_f = n_y - 2$. While $\text{RMSE}_k = 0.03$ is observed for the base scheme, we see roundoff RMSE around 7.0×10^{-8} for the CD case and 3.0×10^{-7} for the projection case. Furthermore, Proj requires an increased processing time by about 20% compared to the CD case. Therefore, we use the CD scheme rather than the projection approach to provide zero-divergence measurement data for data-assimilation.

IX. DATA-ASSIMILATION NUMERICAL EXAMPLE

We consider the grid of Section VIII to study MHD data assimilation. Also, we set $n_{x_{L0}} = 14$, $n_{x_{Lf}} = 19$, $n_{y_{L0}} = 22$, and $n_{y_{Lf}} = 42$. We assume that we have the observation model (4.1), where $C_k = \begin{bmatrix} C^{[17,30]} \\ C^{[17,34]} \end{bmatrix}$, and $C^{[i,j]} \triangleq \begin{bmatrix} 0_{6 \times 6(n_y(i-1)+(j-1))} & I_{6 \times 6} & 0_{6 \times 6(n_y(i+1)+(n_y-j))} \end{bmatrix} \in \mathbb{R}^{6 \times n}$, that is, measurements are obtained only from cells [17, 30] and [17, 34]. We set $R_k = 10^{-6} I_{m \times m}$. Both measured data from base scheme (Base) and central difference scheme (CD) are investigated; see Section VIII. Then, we implement LUKF and, to enforce the zero-divergence constraint (4.2), we implement LECUKF and PLUKF. Also, we set as initial conditions $\hat{x}_{0|0} = \hat{x}_{0|0}^p = x_0 + \delta x_0$ and $P_{L,0|0}^{xx} = P_{L,0|0}^{xpp} = 0.001 I_{n \times n}$, where $x_0 \in \mathbb{R}^n$ is the true initial value used in Section VIII and δx_0 is a zero-mean Gaussian random vector with covariance $0.0004 I_{n \times n}$.

Figure 3 shows the y components of magnetic field at $kT = 15$ s for eight different combinations of data (Base or CD), data-assimilation algorithm (LUKF, LECUKF, and PLUKF) and process model (Base or CD), namely, (i) data from base (Base) scheme with LUKF algorithm using Base model - Base + LUKF, (ii) Base data and LECUKF algorithm using Base model - Base + LECUKF, (iii) CD data and

LUKF algorithm using Base model - CD + LUKF, (iv) CD data and LUKF algorithm using CD model - CD + LUKF (CD), (v) CD data and LECUKF algorithm using Base model - CD + LECUKF, (vi) CD data and LECUKF algorithm using CD model - CD + LECUKF (CD), (vii) CD data and PLUKF algorithm using Base model - CD + PLUKF, and (viii) CD data and PLUKF algorithm using CD model - CD + PLUKF (CD). Though a bowshock wave is created in all cases, the results are substantially different.

Figure 4a shows RMSE_k as a function of time for $i_0 = n_{x_{L0}}$, $i_f = n_{x_{Lf}}$, $j_0 = n_{y_{L0}}$, $j_f = n_{y_{Lf}}$. While $\text{RMSE}_k = 0.04$ is observed for the three cases where LUKF is used, $\text{RMSE}_k = 7 \times 10^{-5}$ is observed for the cases where LECUKF is employed to enforce the zero-divergence constraint, and $\text{RMSE}_k = 10^{-7}$ is obtained for the cases where PLUKF is used. Figure 4b shows RMSE_k for $i_0 = 3$, $i_f = n_x - 2$, $j_0 = 3$, $j_f = n_y - 2$, that is, for the full grid. In this case, LECUKF yields RMSE_k values similar to those obtained when LUKF is used, whereas PLUKF guarantees zero divergence within numerical precision. Note that, although PLUKF provides magnetic-field estimates satisfying the zero-divergence constraint with the smallest RMSE, we cannot claim the same about the accuracy of its estimates because we do not have a closed solution of the MHD equations for the example under investigation to compare our results with. Since LECUKF and PLUKF are suboptimal estimators, they can provide less accurate estimates than does LUKF.

It is important to mention that LECUKF and PLUKF has processing time similar to the processing time of LUKF. Note that, even if the CD model is used during forecast and CD data is used during data assimilation, LUKF does not produce data-assimilation estimates that satisfy the zero-divergence constraint because the Kalman gain given by (5.10) does not take the constraint into account.

X. CONCLUDING REMARKS

We investigated the problem of enforcing the zero-divergence constraint in both the simulation and the data assimilation of magnetohydrodynamics. This was accomplished by using either a field-interpolated central difference or a projection scheme for simulation. For data assimilation, we investigated the localized unscented Kalman filter and, to enforce the zero-divergence constraint, we presented the localized equality-constrained unscented Kalman filter and the projected localized unscented Kalman filter. A two-dimensional example illustrated the problems above.

Results show that, even if zero-divergence model and measured data are used for data-assimilation, the LUKF estimates do not satisfy the zero-divergence property. However, whenever LECUKF is used to enforce the zero-divergence constraint, the state estimates in the localized region satisfy this equality constraint within numerical precision. PLUKF enforces the zero-divergence constraint in the full grid but without recursively feeding the zero-divergence estimates back in the next estimation step.

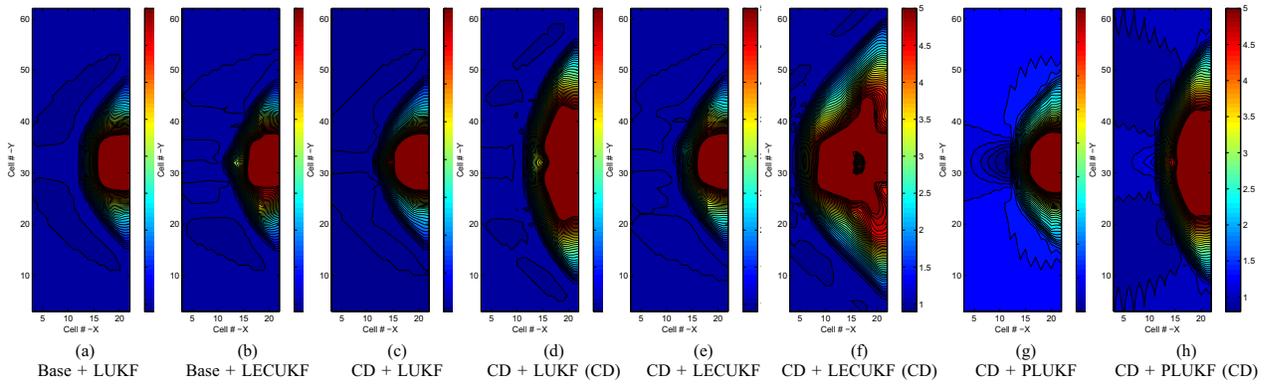


Fig. 3. B_y components of magnetic field at $t = 15$ s obtained from two-dimensional MHD data assimilation by algorithms LUKF, LECUKF, and PLUKF using data obtained from the base (Base) and central difference (CD) approaches.

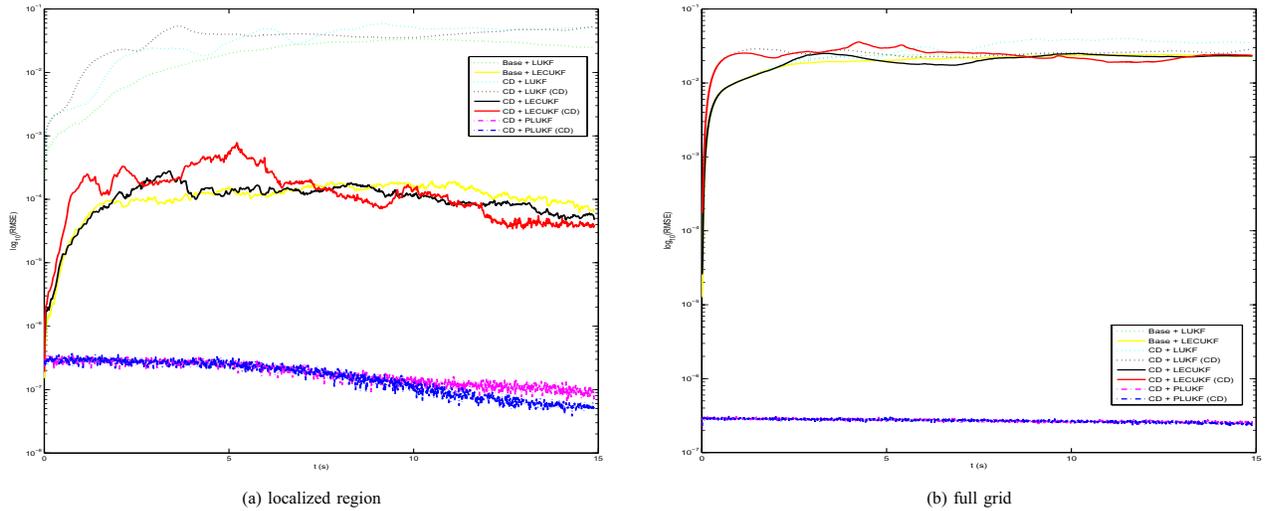


Fig. 4. RMSE index (8.1) evolving with time $t = kT$ for estimated data for six different combinations of data (Base or CD), data-assimilation algorithm (LUKF, LECUKF, PLUKF) plus model (Base or CD).

As future work, we suggest the investigation of an example whose closed solution is known such that the tradeoff between the accuracy of the estimates provided by LUKF, LECUKF, and PLUKF and the accuracy of the satisfaction of the zero-divergence constraint can be evaluated.

ACKNOWLEDGEMENTS

We thank Gábor Tóth for helpful discussions and Jaganath Chandrasekhar for assistance with the MHD code.

REFERENCES

- [1] G. J. Bierman. *Factorization Methods for Discrete Sequential Estimation*. New York: Academic, 1977.
- [2] J. U. Brackbill and D. C. Barnes. The Effect of Nonzero $\nabla \cdot B$ on the Numerical Solution of the Magnetohydrodynamic Equations. *Journal of Computational Physics*, 35:426–430, 2002.
- [3] J. Chandrasekar, A. J. Ridley, and D. S. Bernstein. A comparison of the extended and unscented Kalman filters for discrete-time systems with nondifferentiable dynamics. In *Proceedings of the American Control Conference*, New York City, NY, USA, pp. 4431–4436, 2007.
- [4] J. Chandrasekar, I. S. Kim, A. J. Ridley and D. S. Bernstein. Reduced-Order Covariance-Based Unscented Kalman Filtering with Complementary Steady-State Correlation. In *Proceedings of the American Control Conference*, New York City, NY, USA, pp. 4452–4457, 2007.

- [5] F. E. Daum. Nonlinear filters: Beyond the Kalman filter. *IEEE Aerospace and Electronics Systems Magazine*, 20(8):57–69, 2005.
- [6] J. P. Friedberg. *Ideal Magnetohydrodynamics*. Plenum Press, 1987.
- [7] C. Groth, D. D. Zeeuw, T. Gombosi, and K. Powell. Global 3D MHD Simulation of a Space Weather Event: CME Formation, Interplanetary Propagation, and Interaction with the Magnetosphere. *Journal of Geophysical Research*, 105:25053–25078, 2000.
- [8] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, Inc., New York – NY, USA, 1970; reprinted by Dover, 2007.
- [9] S. J. Julier and J. K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92:401–422, March 2004.
- [10] R. J. Leveque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2002.
- [11] B. O. S. Teixeira, J. Chandrasekar, L. A. B. Tôres, L. A. Aguirre, and D. S. Bernstein. Unscented Filtering for Equality-Constrained Nonlinear Systems. In *Proceedings of the American Control Conference*, Seattle, WA, June, 2008.
- [12] B. O. S. Teixeira, J. Chandrasekar, L. A. B. Tôres, L. A. Aguirre, and D. S. Bernstein. State estimation for linear and nonlinear equality-constrained systems. *International Journal of Control*, submitted.
- [13] G. Tóth. The $\nabla \cdot B = 0$ Constraint in Shock-Capturing Magnetohydrodynamics Codes. *Journal of Computational Physics*, 161:605–652, 2000.