Proceedings of the 2007 American Control Conference
Marriott Marquis Hotel at Times Square
New York City, USA, July 11-13, 2007

Reduced-Order Covariance-Based Unscented Kalman Filtering with Complementary Steady-State Correlation

J. Chandrasekar, I.S. Kim, A. J. Ridley, and D. S. Bernstein

Abstract— **In this paper, we discuss an extension of the unscented Kalman filter that propagates a surrogate reduced-order covariance and also uses a complementary static estimator gain based on the steady-state correlation between the error in the estimates of the state and measurements to obtain estimates of the entire state.**

I. Introduction

State estimation for very large scale systems remains an area of interest research. These systems arise in applications based on spatially distributed models or spatially discretized partial differential equations. Weather forecasting and related atmospheric applications are the main driver for this line of research [1, 2]. Although the literature on reduced-order filtering extends back several decades [3, 4], the challenge in addressing very large scale systems is to propagate the covariance efficiently, especially in view of the fact that covariance propagation is $O(n^3)$ in computational complexity, where n is the number of states.

To address the problem of computational complexity, a reduced-order error-covariance propagation algorithm is developed in [5] based on balanced reduction, and this algorithm is compared to several alternative reduced-order error-covariance propagation algorithms in [6].

In the present paper we extend one of the approaches considered in [6] to nonlinear systems by using the unscented Kalman filter [7]. Since balancing is usually not feasible for systems of very large order, we consider nonlinear extensions of only the algorithms studied in [6] that avoid the need for balancing. These algorithms include the localized unscented Kalman filter (LUKF), which is essentially an unscented Kalman filter applied to a truncated model that includes all states that affect the measurements, as well as LUKF augmented by static estimator gains that are based on complementary steady-state error correlations.

To compare the performance of theestimation algorithms, we consider three examples that are computationally tractable on single-processor machines. First, we consider a finite-volume compressible hydrodynamic simulation for one-dimensional. Next, we consider a two-dimensional finite-volume magnetohydrodynamic (MHD) simulation using the BATSRUS MHD code developed in [8]. Finally, we consider a one-dimensional model of the Global Ionosphere-Thermosphere Model (GITM) [11], which considers the effect of solar flux on the dynamics of the atmosphere.

This research was supported by the National Science Foundation, through Grant ATM-0325332 and CNS-0539053 to the University of Michigan, Ann Arbor, USA. The authors are with the University of Michigan, Ann Arbor, MI-48109, dsbaero@umich.edu.

II. The Unscented Kalman Filter

Consider the discrete-time nonlinear system with dynamics

$$x_{k+1} = f(x_k, u_k, k) + w_k \tag{2.1}$$

and measurements

$$y_k = h(x_k, k) + v_k, \tag{2.2}$$

where $x_k, w_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$, and $y_k, v_k \in \mathbb{R}^p$. The input u_k and output y_k are assumed to be measured, and w_k and v_k are uncorrelated zero-mean white noise processes with covariances Q_k and R_k, respectively. We assume that R_k is positive definite. If the dynamics and the measurement map are linear, the Kalman filter yields the optimal (minimum variance) estimates of the state x_k. The Kalman filter depends on the error covariance which is propagated using the Riccati equation [9].

In this paper, we consider the unscented Kalman filter (UKF) [7]. Assume that $\overline{x} \in \mathbb{R}^n$, $\overline{P} \in \mathbb{R}^{n\times n}$ is positive semidefinite and $\lambda > 0$. The unscented transformation is used to obtain $2n+1$ sample points $X_i \in \mathbb{R}^n$ and corresponding weights $\gamma_{\mathrm{x},i}$ and $\gamma_{\mathrm{P},i}$, for $i = 0, \ldots, 2n$, so that the weighted mean and the weighted variance of the sample points is $\overline{x}$ and $\overline{P}$, respectively. The unscented transformation $X = \Psi(\overline{x}, \overline{P}, \lambda)$ of $\overline{x}$ with covariance $\overline{P}$ is defined by

$$X_i = \begin{cases} \overline{x}, & \text{if } i = 0, \\ \overline{x} + \tilde{P}_i, & \text{if } i = 1, \ldots, n, \\ \overline{x} - \tilde{P}_{i-n}, & \text{if } i = n+1, \ldots, 2n, \end{cases} \tag{2.3}$$

where $\tilde{P} \triangleq \left(\lambda \overline{P}\right)^{1/2}$, $\tilde{P}_i$ is the ith column of $\tilde{P} \in \mathbb{R}^{n\times n}$, $X \in \mathbb{R}^{n\times 2n+1}$ has entries $X = \begin{bmatrix} X_0 & \cdots & X_{2n} \end{bmatrix}$ and $\lambda > 0$ determines the spread of the sample points around $\overline{x}$. Note that $\sum_{i=0}^{2n} \gamma_{\mathrm{x},i} X_i = \overline{x}$ and $\sum_{i=0}^{2n} \gamma_{\mathrm{P},i}(X_i - \overline{x})(X_i - \overline{x})^{\mathrm{T}} = \overline{P}$, where the weights are defined by $\gamma_{\mathrm{x},0} \triangleq 1 - \frac{n}{\lambda}$, $\gamma_{\mathrm{P},0} \triangleq 1 - \frac{n}{\lambda} + (1 - \frac{\lambda}{n} + \beta)$, and for $i = 1, \ldots, 2n$, $\gamma_{\mathrm{x},i} = \gamma_{\mathrm{P},i} \triangleq \frac{1}{2\lambda}$.

UKF involves simulating $2n+1$ copies of the model and using these ensembles to approximate the mean and error covariance. We assume that an initial estimate x_0^{f} of the state x_0 is given, and the covariance of error in the initial condition is $P_0^{\mathrm{f}} \in \mathbb{R}^{n\times n}$.

For all $k \geqslant 0$, the analysis step of UKF is given by

$$x_k^{\mathrm{da}} = x_k^{\mathrm{f}} + K_k(y_k - y_k^{\mathrm{f}}), \tag{2.4}$$

$$y_k^{\mathrm{f}} = h(x_k^{\mathrm{f}}, k), \tag{2.5}$$

$$X_k^{\mathrm{da}} = \Psi(x_k^{\mathrm{da}}, P_k^{\mathrm{da}}, \lambda), \tag{2.6}$$

$$P_k^{\mathrm{da}} = P_k^{\mathrm{f}} - K_k P_{\mathrm{yy},k} K_k^{\mathrm{T}}, \tag{2.7}$$

where

$$K_k = P_{\mathrm{xy},k} P_{\mathrm{yy},k}^{-1}, \tag{2.8}$$

$$P_{\mathrm{xy},k} = \sum_{i=0}^{2n} \gamma_{\mathrm{P},i} (X_{i,k}^{\mathrm{f}} - x_k^{\mathrm{f}})(Y_{i,k}^{\mathrm{f}} - y_k^{\mathrm{f}})^{\mathrm{T}}, \tag{2.9}$$

$$P_{\mathrm{yy},k} = \sum_{i=0}^{2n} \gamma_{\mathrm{P},i} (Y_{i,k}^{\mathrm{f}} - y_k^{\mathrm{f}})(Y_{i,k}^{\mathrm{f}} - y_k^{\mathrm{f}})^{\mathrm{T}} + R_k, \tag{2.10}$$

$$X_k^{\mathrm{f}} = \Psi(x_k^{\mathrm{f}}, P_k^{\mathrm{f}}, \lambda), \tag{2.11}$$

$$Y_{i,k}^{\mathrm{f}} = h(X_{i,k}^{\mathrm{f}}, k), \tag{2.12}$$

and the forecast step of UKF is given by

$$\hat{X}_{i,k+1}^{\mathrm{f}} = f(X_{i,k}^{\mathrm{da}}, u_k, k), \quad x_{k+1}^{\mathrm{f}} = \sum_{i=0}^{2n} \gamma_{\mathrm{x},i} \hat{X}_{i,k+1}^{\mathrm{f}}, \tag{2.13}$$

$$P_{k+1}^{\mathrm{f}} = \sum_{i=0}^{2n} \gamma_{\mathrm{P},i} (\hat{X}_{i,k+1}^{\mathrm{f}} - x_{k+1}^{\mathrm{f}})(\hat{X}_{i,k+1}^{\mathrm{f}} - x_{k+1}^{\mathrm{f}})^{\mathrm{T}} + Q_k. \tag{2.14}$$

Since UKF involves $2n + 1$ model updates given by (2.13), the computational burden of UKF is of the order $(2n + 1)n^2 = 2n^3 + n^2$. If n is large, for example, in finite volume discretization of partial differential equations, then the computational burden of implementing UKF is enormous. Next, we consider an extension of UKF that approximates the error covariance corresponding to only a specific subspace of the state, thereby reducing the number of ensembles needed.

III. Localized Unscented Kalman Filter (LUKF)

Assume that the state $x_k \in \mathbb{R}^n$ has components

$$x_k = \begin{bmatrix} x_{\mathrm{L},k}^{\mathrm{T}} & x_{\mathrm{E},k}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}, \tag{3.1}$$

where $x_{\mathrm{L},k} \in \mathbb{R}^{n_{\mathrm{L}}}$ and $x_{\mathrm{E},k} \in \mathbb{R}^{n_{\mathrm{E}}}$. Also, assume that the measurements depend on the state x_{L} so that y_k can be expressed as

$$y_k = h(x_{\mathrm{L},k}, k) + v_k. \tag{3.2}$$

The objective is to directly inject the measurement data y_k into only the states corresponding to the estimate of $x_{\mathrm{L},k}$ by using a reduced-order surrogate error covariance. For example, in weather prediction models involving spatial dimensions, $x_{\mathrm{L},k}$ may represent the states corresponding to a small region surrounding the location where measurements are available, and $x_{\mathrm{E},k}$ may represent the states that are outside this localized region.

Assume that for all $k \geqslant 0$, the error covariance P_k^{f} of UKF has the structure

$$P_k^{\mathrm{f}} = \begin{bmatrix} P_{\mathrm{L},k}^{\mathrm{f}} & 0 \\ 0 & 0 \end{bmatrix}, \tag{3.3}$$

where $P_{\mathrm{L},k}^{\mathrm{f}} \in \mathbb{R}^{n_{\mathrm{L}} \times n_{\mathrm{L}}}$ represents the covariance of error corresponding to the state $x_{\mathrm{L},k}$. Hence, it follows from (2.3) and (3.3) that if $X_k^{\mathrm{f}} = \Psi(x_k^{\mathrm{f}}, P_k^{\mathrm{f}}, \lambda)$ then for $i = n_{\mathrm{L}} + 1, \ldots, n, n + n_{\mathrm{L}} + 1, \ldots, 2n$, $X_{i,k}^{\mathrm{f}} = X_{0,k}^{\mathrm{f}} = x_k^{\mathrm{f}}$. Since $2n_{\mathrm{E}} + 1$ ensembles are exactly the same, it suffices to retain only one such ensemble. Therefore, the number of ensembles required is reduced from $2n + 1$ to $2n_{\mathrm{L}} + 1$. Furthermore, it follows from (3.3) that instead of a $n \times n$ error covariance, only a $n_{\mathrm{L}} \times n_{\mathrm{L}}$ reduced-order error covariance has to be estimated using the $2n_{\mathrm{L}} + 1$ ensembles.

The data assimilation step of LUKF is given by

$$x_{\mathrm{L},k}^{\mathrm{da}} = x_{\mathrm{L},k}^{\mathrm{f}} + K_{\mathrm{L},k}(y_k - y_k^{\mathrm{f}}), \quad x_{\mathrm{E},k}^{\mathrm{da}} = x_{\mathrm{E},k}^{\mathrm{f}} \tag{3.4}$$

$$y_k^{\mathrm{f}} = h(x_{\mathrm{L},k}^{\mathrm{f}}, k), \tag{3.5}$$

$$X_{\mathrm{L},k}^{\mathrm{da}} = \Psi(x_{\mathrm{L},k}^{\mathrm{da}}, P_{\mathrm{L},k}^{\mathrm{da}}, \lambda), \tag{3.6}$$

$$P_{\mathrm{L},k}^{\mathrm{da}} = P_{\mathrm{L},k}^{\mathrm{f}} - K_{\mathrm{L},k} P_{\mathrm{yy},k} K_{\mathrm{L},k}^{\mathrm{T}}, \tag{3.7}$$

where

$$K_{\mathrm{L},k} = P_{\mathrm{x_L y},k} P_{\mathrm{yy},k}^{-1}, \tag{3.8}$$

$P_{\mathrm{yy},k}$ is given by (2.10), $Y_{i,k}$ is given by (2.12), $P_{\mathrm{x_L y},k}$ and $X_{\mathrm{L},k}^{\mathrm{f}}$ are given by (2.9) and (2.11), respectively, with $X_{i,k}^{\mathrm{f}}$ replaced by $X_{\mathrm{L},i,k}^{\mathrm{f}}$, x_k^{f} replaced by $x_{\mathrm{L},k}^{\mathrm{f}}$, and P_k^{f} replaced by $P_{\mathrm{L},k}^{\mathrm{f}}$, and for $i = 0, \ldots, 2n_{\mathrm{L}}$, $X_{\mathrm{L},i,k}^{\mathrm{f}} \in \mathbb{R}^{n_{\mathrm{L}}}$ is the $(i+1)$th column of $X_{\mathrm{L},k}^{\mathrm{f}}$.

Next, for all $i = 0, \ldots, 2n_{\mathrm{L}}$, define $X_{i,k}^{\mathrm{da}} \in \mathbb{R}^n$ by $X_{i,k}^{\mathrm{da}} \triangleq \begin{bmatrix} \left(X_{\mathrm{L},i,k}^{\mathrm{da}}\right)^{\mathrm{T}} & \left(x_{\mathrm{E},k}^{\mathrm{da}}\right)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$, where $X_{\mathrm{L},i,k}^{\mathrm{da}} \in \mathbb{R}^{n_{\mathrm{L}}}$ is the $(i+1)$th column of $X_{\mathrm{L},k}^{\mathrm{da}}$. It follows from (3.3) that the correlations corresponding to the error in the state $x_{\mathrm{E},k}$ are assumed to be zero, and therefore, the estimate $x_{\mathrm{E},k}^{\mathrm{da}}$ of the state $x_{\mathrm{E},k}$ in all the ensembles $X_{i,k}^{\mathrm{da}}$ of LUKF are the same. However, the estimate of the state $x_{\mathrm{L},k}$ is different in each ensemble. The forecast step of LUKF is given by

$$\hat{X}_{i,k+1}^{\mathrm{f}} = f(X_{i,k}^{\mathrm{da}}, u_k, k), \; x_{k+1}^{\mathrm{f}} = \sum_{i=0}^{2n_{\mathrm{L}}} \gamma_{\mathrm{x},i} \hat{X}_{i,k+1}^{\mathrm{f}}. \tag{3.9}$$

Next, for $i = 0, \ldots, 2n_{\mathrm{L}}$, let $\hat{X}_{i,k}^{\mathrm{f}} \in \mathbb{R}^n$ have entries

$$\hat{X}_{i,k}^{\mathrm{f}} = \begin{bmatrix} \left(\hat{X}_{\mathrm{L},i,k}^{\mathrm{f}}\right)^{\mathrm{T}} & \left(\hat{X}_{\mathrm{E},i,k}^{\mathrm{f}}\right)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}, \tag{3.10}$$

with $\hat{X}_{\mathrm{L},i,k}^{\mathrm{f}} \in \mathbb{R}^{n_{\mathrm{L}}}$ and $\hat{X}_{\mathrm{E},i,k}^{\mathrm{f}} \in \mathbb{R}^{n_{\mathrm{E}}}$. Finally, to account for the increase in the error covariance due to the process noise, represented by $Q_{\mathrm{L},k}$, the surrogate covariance of the error in the estimate of $x_{\mathrm{L},k}$ is given by

$$P_{\mathrm{L},k+1}^{\mathrm{f}} = \sum_{i=0}^{2n_{\mathrm{L}}} \gamma_{\mathrm{P},i} (\hat{X}_{\mathrm{L},i,k+1}^{\mathrm{f}} - x_{\mathrm{L},k+1}^{\mathrm{f}})(\hat{X}_{\mathrm{L},i,k+1}^{\mathrm{f}} - x_{\mathrm{L},k+1}^{\mathrm{f}})^{\mathrm{T}} + Q_{\mathrm{L},k} \tag{3.11}$$

LUKF involves $2n_{\mathrm{L}} + 1$ model updates and therefore the number of computations involved is of the order $(2n_{\mathrm{L}}+1)n^2$. Hence, if $n_{\mathrm{L}} \ll n$, then LUKF is computationally efficient compared to UKF.

IV. Complementary Steady-State Correlation

Since ignoring the correlation between the error in the estimates of the states $x_{\mathrm{L},k}$ and $x_{\mathrm{E},k}$ in LUKF may result in poor estimates, we consider a modification of LUKF that uses a constant correlation between the error in the estimates of the states $x_{\mathrm{L},k}$ and $x_{\mathrm{E},k}$. In the following sections, we assume that $Q_k = Q$ and $R_k = R$ for all $k \geqslant 0$.

If the dynamics and the measurement map in (2.1) and (2.2) are linear and time-invariant, then, the error covariance

generally converges to a steady-state value. If the dynamics are nonlinear, then there is no guarantee that UKF or LUKF will reach a statistical steady-state. However, simulations may indicate that after a certain period of time, the performance of the estimators do not vary significantly, and in that case, we assume that the estimator has almost reached statistical steady-state.

A. LUKF with Complementary Open-Loop Correlation (LUKFCOLC)

If the dynamics are linear and time-invariant, that is $f(x,u,k) = Ax + Bu$ and $h(x,k) = Cx$ for all $k \geqslant 0$, and (A, Q) is stabilizable, then the steady-state state covariance $P_{\rm xx}$ is the solution of a Lyapunov equation. Furthermore, the steady state correlation $P_{\rm xy}$ between the measurement y_k and the state x_k is given by $P_{\rm xy} = P_{\rm xx}C^{\rm T}$.

However, since the dynamics are nonlinear, we approximate the steady-state state covariance by using Monte Carlo simulations. Consider N copies of the open-loop model of the system (2.1)-(2.2) so that for $i = 1, \ldots, N$,

$$\tilde{x}_{i,k+1} = f(\tilde{x}_{i,k}, u_k, k) + \tilde{w}_{i,k}, \;\; \tilde{y}_{i,k} = h(\tilde{x}_{i,k}, k) + \tilde{v}_{i,k}, \quad (4.1)$$

where $\tilde{x}_{i,0}$ is a random variable with the specified mean x_0 and variance $P_0^{\rm f}$, and $\tilde{w}_{i,k}$ and $\tilde{v}_{i,k}$ are sampled from zero-mean white processes with variances Q and R, respectively. Next, we define an approximation of the steady state open-loop correlation $P_{\rm OL,xy}$ and $P_{\rm OL,yy}$ by

$$P_{\rm OL,xy} \triangleq \lim_{k\to\infty} \frac{1}{N-1}\sum_{i=1}^{N}(\tilde{x}_{i,k} - \overline{x}_k)(\tilde{y}_{i,k} - \overline{y}_k)^{\rm T}, \quad (4.2)$$

$$P_{\rm OL,yy} \triangleq \lim_{k\to\infty} \frac{1}{N-1}\sum_{i=1}^{N}(\tilde{y}_{i,k} - \overline{y}_k)(\tilde{y}_{i,k} - \overline{y}_k)^{\rm T}, \quad (4.3)$$

where $\overline{x}_k \triangleq \frac{1}{N}\sum_{i=1}^{N}\tilde{x}_{i,k}$, $\overline{y}_k \triangleq \frac{1}{N}\sum_{i=1}^{N}\tilde{y}_{i,k}$.

Alternatively, the unscented transformation can also be used to approximate the steady state open-loop state covariance. Note that the state covariance of (2.1) is the same as the open-loop error covariance, that is the covariance of error of an estimator when the estimator gain is zero. Hence, we use (2.4)-(2.14) with $K_k = 0$ for all $k \geqslant 0$, and define $P_{\rm OL,xy}$ and $P_{\rm OL,yy}$ by $P_{\rm OL,xx} \triangleq \lim_{k\to\infty} P_{\rm xy,k}$, $P_{\rm OL,yy} \triangleq \lim_{k\to\infty} P_{\rm yy,k}$. If n is small, then the computational burden of using the open-loop unscented Kalman filter to estimate the open-loop error correlation is small. However, when n is large, approximating the error covariance by using Monte Carlo simulations with a small N is computationally more efficient.

Finally, we define the static estimator gain $K_{\rm OL} \in \mathbb{R}^{n\times p}$ based on the steady-state open-loop correlations by $K_{\rm OL} \triangleq P_{\rm OL,xy}P_{\rm OL,yy}^{-1}$ and let $K_{\rm OL}$ have entries

$$K_{\rm OL} = \begin{bmatrix} (K_{\rm OL,L})^{\rm T} & (K_{\rm OL,E})^{\rm T} \end{bmatrix}^{\rm T}, \quad (4.4)$$

where $K_{\rm OL,L} \in \mathbb{R}^{n_{\rm L}\times p}$ and $K_{\rm OL,E} \in \mathbb{R}^{n_{\rm E}\times p}$. The forecast step of LUKFCOLC is given by (3.9) - (3.11). The analysis step of the LUKFCOLC is given by

$$x_{{\rm L},k}^{\rm da} = x_{{\rm L},k}^{\rm f} + K_{{\rm L},k}(y_k - y_k^{\rm f}), \quad (4.5)$$

$$x_{{\rm E},k}^{\rm da} = x_{{\rm E},k}^{\rm f} + K_{\rm OL,E}(y_k - y_k^{\rm f}), \quad (4.6)$$

$$y_k^{\rm f} = h(x_{{\rm L},k}^{\rm f}, k), \quad (4.7)$$

$$X_{{\rm L},k}^{\rm da} = \Psi(x_{{\rm L},k}^{\rm da}, P_{{\rm L},k}^{\rm da}, \lambda), \quad (4.8)$$

$$P_{{\rm L},k}^{\rm da} = P_{{\rm L},k}^{\rm f} - K_{{\rm L},k}P_{{\rm yy},k}K_{{\rm L},k}^{\rm T}, \quad (4.9)$$

where $P_{{\rm yy},k}$ and $K_{{\rm L},k}$ are defined in (2.10) and (3.8).

Note that injecting measurement data in an estimator affects the error covariances and hence, the actual closed-loop error correlation between y_k and the error in estimates $x_k^{\rm f} - x_k$ will be different from the open-loop error correlation. However, (4.6) implies that the estimator gain corresponding to the estimate $x_{{\rm E},k}^{\rm da}$ is based on only the open-loop error correlation and is not aware of the change in correlation due to data injection.

B. LUKF with Complementary Closed-Loop Correlation (LUKFCCLC)

Next, we use a static estimator gain that is based on the closed-loop steady-state correlations. Specifically, we estimate the steady-state correlations between the error in the estimates when LUKF is used for state estimation. We assume that LUKF has reached a statistical steady-state when the performance of LUKF does not change significantly.

The Monte-Carlo procedure to determine the steady-state closed-loop correlation is as follows. First, we simulate N copies of the open-loop model of the system as shown in (4.1) and obtain outputs $\tilde{y}_{i,k}$. Next, for $i = 1, \ldots, N$, we perform state estimation using LUKF with the outputs $\tilde{y}_{i,k}$. Let $\tilde{y}_{i,k}^{\rm f}$ and $\tilde{x}_{i,k}^{\rm f}$ be the estimate of $\tilde{y}_{i,k}$ and $\tilde{x}_{i,k}$, respectively, provided by the ith simulation of LUKF. We approximate the steady-state closed-loop correlations by (4.2) and (4.3) with $\tilde{x}_{i,k} - \overline{x}_k$ replaced by $\tilde{x}_{i,k}^{\rm f} - \tilde{x}_{i,k}$, and $\tilde{y}_{i,k} - \overline{y}_k$ replaced by $\tilde{y}_{i,k}^{\rm f} - \tilde{y}_{i,k}$. Note that $\tilde{y}_{i,k}$, $\tilde{x}_{i,k}$, $\tilde{y}_{i,k}^{\rm f}$, and $\tilde{x}_{i,k}^{\rm f}$ are all simulation outputs and hence $P_{\rm CL,xy}$ and $P_{\rm CL,yy}$ can be evaluated.

Alternatively, the unscented transformation can also be used to obtain an estimate of the closed-loop error correlations. To do this, we first use LUKF with one of the simulated measurement data, for example, $\tilde{y}_{1,k}$, to obtain estimates $\tilde{x}_{1,k}^{\rm f}$ of the state $\tilde{x}_{1,k}$ for $k \geqslant 0$. Assuming $K_{\rm L,k}$ does not vary significantly after a sufficiently long time interval, we define the steady-state LUKF estimator gain $K_{\rm L}$ by $K_{\rm L} \triangleq \lim_{k\to\infty} K_{{\rm L},k}$, where $K_{{\rm L},k}$ is the estimator gain given by (3.8) when obtaining the estimate $\tilde{x}_{1,k}^{\rm f}$. Note that LUKF ignores correlations between certain states and hence cannot be used to estimate the closed-loop error correlation. Instead, we use the unscented transformation to estimate the closed-loop steady-state error correlations. Specifically, we use (2.4)-(2.14) with $K_k = \begin{bmatrix} (K_{\rm L})^{\rm T} & 0 \end{bmatrix}^{\rm T}$, for all $k \geqslant 0$, and view the correlations $P_{{\rm xy},k}$ and $P_{{\rm yy},k}$ in (2.9) and (2.10) as an estimate of the closed-loop error correlations of LUKF. We then estimate the closed-loop steady-state error correlations $P_{\rm CL,xy}$ and $P_{\rm CL,yy}$ by $P_{\rm CL,xy} = \lim_{k\to\infty} P_{{\rm xy},k}$

and $P_{\mathrm{CL,yy}} = \lim_{k\to\infty} P_{\mathrm{yy},k}$. Finally, the static estimator gain that is based on the steady-state closed-loop error correlations is given by $K_{\mathrm{CL}} = P_{\mathrm{CL,xy}} P_{\mathrm{CL,yy}}^{-1}$, with entries $K_{\mathrm{CL}} = \left[\begin{array}{cc} (K_{\mathrm{CL,L}})^{\mathrm{T}} & (K_{\mathrm{CL,E}})^{\mathrm{T}} \end{array} \right]^{\mathrm{T}}$. The forecast step of LUKFCCLC is given by (3.9) - (3.11), and the analysis step of LUKFCCLC is given by (4.5)-(4.9) with $K_{\mathrm{OL,E}}$ replaced by $K_{\mathrm{CL,E}}$ in (4.6).

V. One-Dimensional Hydrodynamics

First, we consider state estimation of one-dimensional hydrodynamic flow. The flow of an inviscid, compressible fluid along a one-dimensional channel is governed by Euler's equations. A discrete-time model can be obtained by using a finite-volume based spatial and temporal discretization.

Assume that the channel consists of n identical cells. For all $i = 1, \ldots, n$, let $\rho^{[i]}$, $v^{[i]}$, and $p^{[i]}$ be the density, velocity, and pressure in the ith cell, and define $U^{[i]} \in \mathbb{R}^3$ by $U^{[i]} = \left[\begin{array}{ccc} \rho^{[i]} & v^{[i]} & p^{[i]} \end{array} \right]^{\mathrm{T}}$. We discretize Eulers equations and obtain a model that enables us to update the flow variables at the center of each cell.

Next, define the state vector $x \in \mathbb{R}^{3(n-4)}$ by $x \triangleq \left[\begin{array}{ccc} (U_k^{[3]})^{\mathrm{T}} & \cdots & (U_k^{[n-2]})^{\mathrm{T}} \end{array} \right]^{\mathrm{T}}$. Furthermore, we assume Neumann boundary conditions at either ends of the one-dimensional channel. Let $n = 54$ so that $x \in \mathbb{R}^{150}$. The second-order Rusanov scheme yields a nonlinear discrete-time update model of the form (2.1), where $w_k \in \mathbb{R}^{3(n-4)}$ represents unmodeled drivers and is assumed to be zero-mean white Gaussian process noise with covariance matrix $Q \in \mathbb{R}^{3(n-4)\times 3(n-4)}$, so that the flow variables in only the 5th, 15th, 25th, 35th, and 45th cell are directly affected by w_k. We assume that measurement $y_k \in \mathbb{R}^6$ of density, momentum and energy at cells with indices 24 and 26 is available and v_k is zero-mean white Gaussian noise with covariance matrix $R = 0.01 I_{6\times 6}$.

We simulate the truth model with the initial condition $\varrho_0^{[i]} = 1$, $v_0^{[i]} = 0$, and $p_0^{[i]} = 1$ for $i = 1, \ldots, n$ and obtain measurements y_k. The objective is to estimate the density, momentum, and energy at the cells where measurements of flow variables are unavailable using UKF, LUKF, LUKFCOLC, and LUKFCCLC.

The square root of the sum of the square of the error in the estimates of the energy at cells $1, \ldots, 50$, when measurements y_k are used in the UKF is shown in Figure 1. The error in energy estimates when no data assimilation is performed is also shown in the same figure for comparison. Next, we compare the performance of LUKF for various local grid sizes, that is, we set

$$x_{\mathrm{L}} \triangleq \left[\begin{array}{ccc} (U_k^{[L_1]})^{\mathrm{T}} & \cdots & (U_k^{[L_n]})^{\mathrm{T}} \end{array} \right]^{\mathrm{T}}, \qquad (5.1)$$

where $(L_1, L_n) \in \{(20, 30), (16, 34), (12, 38)\}$. The square root of the sum of the square of the error in energy estimates of LUKF is shown in Figure 1 for the three different local grid sizes. For all three local grid sizes, the performance of UKF is much better than the performance of LUKF because

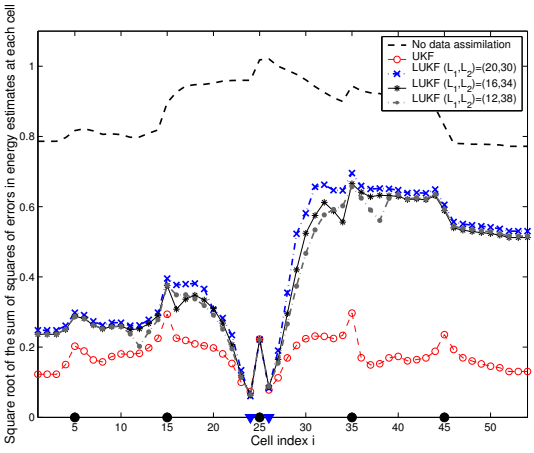

Fig. 1. Plot of the square root of the sum of the square of the error in energy estimates at the various cells using UKF and LUKF with 3 different local grid sizes. Although the local grid size where data is directly injected increases, the performance of LUKF shows only a minor improvement. The cells where disturbance enters the system are indicated by '•' and the cells where measurements are available are indicated by '▼'.

LUKF ignores correlations between the measurement and the states that are outside the local region.

Finally, we obtain the steady-state open-loop and closed-loop error correlations by using the unscented transformation method. Note that the computational effort of determining the steady-state correlations using the unscented transformation is equivalent to the computational effort of using UKF. However, once the steady-state correlations are determined offline, the computational effort of LUKFCOLC and LUKFCCLC while performing the actual data assimilation is similar to that of LUKF which is significantly lower than the computational effort of UKF.

The square root of the sum of the square of the error in energy estimates is shown in Figure 2. We choose $(L_1, L_n) = (20, 30)$ for LUKF, LUKFCOLC, and LUKFCCLC. The performance of LUKFCCLC is better than the performance of LUKFCOLC because LUKFCCLC accounts for the change in the measurement-error correlation when data is injected during estimation.

VI. Two Dimensional Magnetohydrodynamics Using BATSRUS

BATSRUS (Block Adaptive-Tree Solar-wind Roe-type Upwind Scheme) [8] is a finite volume scheme used to model the interactions between the magnetic field of various planets with the solar wind. The dynamics of the flow variables is governed by Euler's equations and Maxwell's electromagnetic equation.

Consider a 2D spatial grid comprising of 4800 square cells that covers a rectangular region spanning the coordinates $-10 \leqslant x_{\mathrm{c}} \leqslant 10$ and $-30 \leqslant y_{\mathrm{c}} \leqslant 30$. We use BATSRUS to model the dynamics of the flow variables density (ρ), velocity (v_x, v_y), pressure (p) and magnetic field (B_x, B_y) in each cell. We choose initial flow conditions so that the flow is supersonic. We assume floating boundary conditions for all cells along the edges, except for two cells at locations

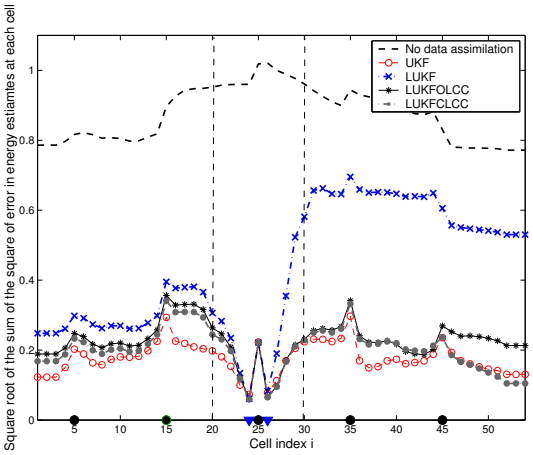

Fig. 2. Square root of the sum of the square of the error in energy estimates from LUKF, LUKFCOLC, and LUKFCCLC. All three estimators use a time varying estimator gain to inject data into the cells with index between 20 and 30. The error in energy estimates from UKF is performed is also plotted for comparison. The performance of LUKFCCLC is close to that of UKF.

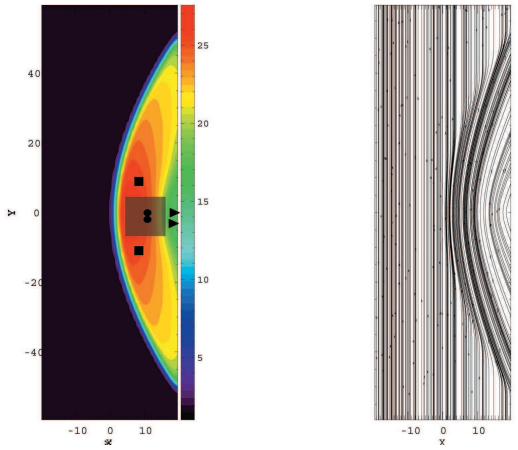

Fig. 3. Magnetic bowshock. The cells where disturbance enters the system is indiacted by '■' and the cells where measurements are available are indicated by '•'. The local region corresponding to the state x_{L} is indicated by the shaded rectangular region around the measurement cells.

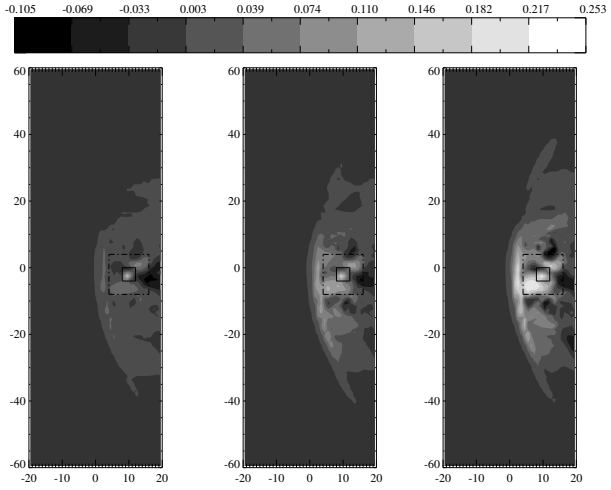

Fig. 4. The difference in the error in the square root of the sum of the square of error in pressure estimates between the no data assimilation case and LUKF (left), LUKFCOLC (middle), and LUKFCCLC (right). The horizontal and vertical axis denote the x and y spatial coordinates. Positive values indicate regions where the estimators improve the estimates of the state compared to the no data assimilation case.

indicated by '▶' in Figure 3 that are assigned reflective boundary conditions so that a bow-shock is created.

A finite-volume discretization of the partial differential equation yields the system dynamics given by (2.1). We assume that w_k in (2.1) is zero-mean white Gaussian process noise that affects only the cells with coordinates indicated by '•' in Figure 3. We assume that noisy measurements y_k of the flow variables at cells indicated by '■' in Figure 3 are available.

The density and magnetic filed lines at $t = 1$ minute are shown in Figure 3. Figure 4 shows a plot of the difference in square root of the sum of the squares of error in energy estimates between the no data assimilation case and LUKF, LUKFCOLC, and LUKFCCLC. Hence, positive values indicate a significant improvement in the estimates. Note that the state dimension $n = 25056$ and since UKF requires $2n + 1 = 50113$ ensembles, we do not use UKF to obtain the state estimates. Also, we use Monte Carlo methods to determine the steady-state correlations used in LUKFCOLC and LUKFCCLC. The local region used in LUKF, LUKFCOLC and LUKFCCLC is shown in Figure 3 by the solid lines and x_{L} contains the state variables in this region.

VII. One Dimensional Global Thermospheric and Ionospheric Model

The Global Ionosphere-Thermosphere Model (GITM) is a 3-dimensional, parallel, spherical code that models the Earth's thermosphere and ionosphere system, which has an altitude range from about 100 km to 1000 km, using a stretched altitude grid [11]. Inputs to GITM include solar ultraviolet (UV) photons, auroral energetic particles, electric fields, and electric currents. The code explicitly solves for the particle densities of neutral and ion species, and neutral-particle, ion, and electron temperatures. Many of these details are discussed in the classic reference [12].

We apply the different data assimilation techniques to 1D GITM with 50 cells defined along the vertical direction. We consider the solar irradiation represented by $F_{10.7}$ as the only input u_k to GITM. $F_{10.7}$ is the flux at 2800 MHz or 10.7-cm wavelength over the entire solar disk. We summarize the features of 1D GITM for data assimilation in Table I. In addition, the 1D GITM cell structure for UKF, LUKF, and LUKF with open-loop or closed-loop steady-state correlation compensation is shown in Figure 5.

The open-loop and closed-loop steady-state correlations of the data assimilation model are obtained by using the unscented transformation. We compare the performance of UKF, LUKF, LUKFCOLC, and LUKFCCLC for 1D GITM. We compare the estimates of the normalized neutral-particle temperature T_n from the different estimation schemes.

Figure 6 compares the estimates of T_n in cell 40. It can

TABLE I: Summary of 1D GITM for data assimilation.

location	Millstone Hill, MA USA
grid	50 grid cells covering 100 km to 750 km (altitude)
input	$F_{10.7}$
states	logarithm of number density of O, O_2, N_2, $N(^4S)$ vertical velocities of O, O_2, N_2, $N(^4S)$ eastward and northward horizontal velocity of neutrals normalized temperature of neutrals T_n logarithm of number density of $O^+(^4S)$, O_2^+ , NO^+ electron temperature T_e 15 states per cell
outputs	electron temperature T_e number density of electron N_e vertical velocity of ions V_i ions temperature T_i in measurement cells 10, 20, 30, and 40

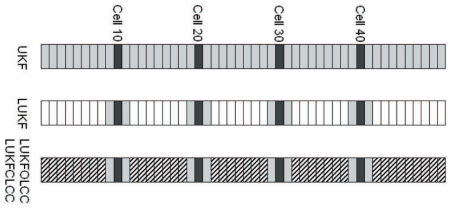

Fig. 5. 1D GITM cell structure for UKF, LUKF, LUKFCOLC and LUKFCCLC. Measurements are taken in cells 10, 20, 30, and 40. LUKF performs data injection only on the local group of 12 cells that include measurement cells. LUKFCOLC and LUKFCCLC consist of LUKF along with data injection into the remaining cells using the static estimator gain.

be seen that LUKFCOLC and LUKFCCLC perform better than LUKF. Furthermore, the estimates from LUKFCCLC are closer to the truth model than LUKFCOLC.

VIII. Conclusion

We present extensions of the the unscented Kalman filter that propagate a reduced-order pseudo error covariance. To compensate for the neglected correlation between certain states and the measurement, we present two methods that use a complementary static estimator gain based on correlations between the measurements and the neglected states. The use of a static estimator gain based on the open-loop and closed-loop correlations helps improve estimation performance without a significant increase in the online computational burden.

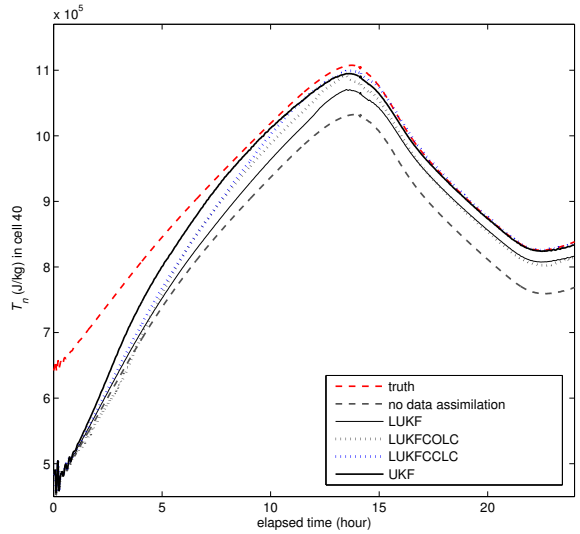

Fig. 6. Comparisons of the normalized neutral-particle temperature in cell 40. Estimates from LUKFCCLC are closer to the truth model than the estimates from LUKFCOLC.

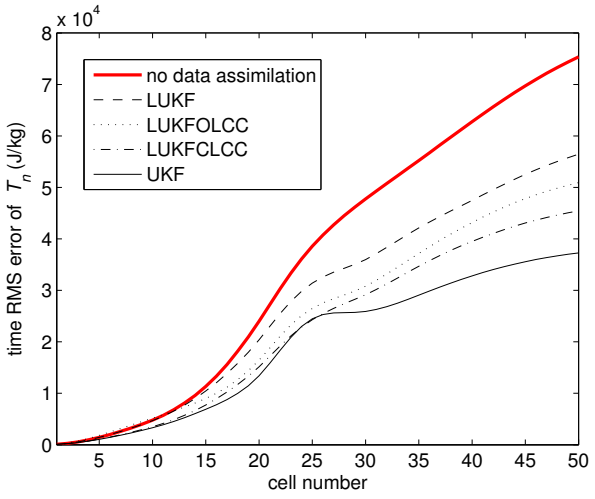

Fig. 7. RMS errors for the normalized neutral-particle temperatures. The errors in estimates from LUKFCCLC is less than the errors in estimates from LUKFCOLC. Moreover, note that the performance of LUKFCOLC and LUKFCCLC is better than the performance of LUKF while all their computational time is almost the same..

References

[1] Y. K. Sasaki and J. S. Goerss, "Satellite Data Assimilation Using NASA Data Systems Test 6 Observations," *Mon. Wea. Rev.*, vol. 110, pp. 16351644, 1982.

[2] J. A. Carton, G. Chepurin, and X. Cao, "A Simple Ocean Data Assimilation Analysis of the Global Upper Ocean 195095. Part I: Methodology," *J. Phy. Ocean,* vol. 30, pp. 294309, 1999.

[3] P. Hippe and C. Wurmthaler, "Optimal Reduced-Order Estimators in the Frequency Domain: The Discrete-Time Case," *Int. J. Contr.*, Vol. 52, pp. 1051-1064, 1990.

[4] W. M. Haddad and D. S. Bernstein, "Optimal Reduced-Order Observer-Estimators," *AIAA J. Guid. Dyn. Contr.,* Vol. 13, pp. 1126-1135, 1990.

[5] B. F. Farrell, and P. J. Ioanannou, "State Estimation Using a Reuced-Order Kalman Filter," *J. Atmo. Sci.*, vol. 58, pp. 3666-3658, 2001.

[6] I. S. Kim, J. Chandrasekar, H. J. Palanthandalam, A. Ridley, and D. S. Bernstein, "State Estimation for Large-Scale Systems Based on Reduced-Order Error-Covariance Propagation," *Amer. Contr. Conf*, New York, 2007.

[7] S. Julier, "The Scaled Unscented Transformation," in *Proc. Amer. Cont. Conf.*, Anchorage, May 2002, pp. 4555 – 4559.

[8] G. Toth, A. Ridley, K. Powell, and T. Gambosi, "A High-Performance Framework for Sun-to-Earth Space Weather Modeling,", *Parallel and Distributed Processing Symposium*, 2005.

[9] B. D. O. Anderson and J. B. Moore, *Optimal Filtering,* Dover Publications Inc., Mineola, NY, 1979.

[10] C. Hirsch, *Numerical Computation of Internal and External Flows*, John Wiley and Sons, 1990.

[11] A. J. Ridley, Y. Deng, and G. Toth, "The Global Ionosphere-Thermosphere Model," *J. Atmos. and Sol.-Terr. Phy.*, vol. 68, pp. 839-864, 2006.

[12] M. H. Rees, *Physics and Chemistry of the Upper Atmosphere,* Cambridge University Press, 1989.