Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Ensemble-On-Demand Kalman Filter for Large-Scale Systems with Time-Sparse Measurements

In Sung Kim, Bruno O. S. Teixeira, and Dennis S. Bernstein

***Abstract*— The ensemble Kalman filter for data assimilation involves the propagation of a collection of ensemble members. Under the assumption of time-sparse measurements, we avoid propagating the ensemble members for all of the time steps by creating an ensemble of models only when a new measurement is made available. We call this algorithm the ensemble-on-demand Kalman filter (EnODKF). We use guidelines for ensemble size within the context of EnODKF, and demonstrate the performance of EnODKF for a representative example, specifically, a heat flow problem.**

I. Introduction

State estimation for spatially distributed systems typically entails nonlinear, high-dimensional dynamics. For these applications, state estimation is known in practice as data assimilation. Applications range from weather forecasting, to oceanography, to structural dynamics [1–3].

Data assimilation methods use variations of the basic formalism of the classical Kalman filter. The most popular methods replace the Riccati equation error covariance propagation of the classical filter with an ensemble of models that approximate the error covariance, which is subsequently used to determine a data injection gain. Two such methods are the ensemble Kalman filter (EnKF) [4], which is based on stochastically sampled drivers, and the unscented Kalman filter (UKF) [5], which is based on deterministically determined drivers for an ensemble of $2n+1$ members, where n is the number of states. In the case of linear systems, UKF exactly reproduces the results of the classical Kalman filter.

Many of the EnKF or UKF applications of interest arise from extremely high-order dynamics. In particular, we are interested in the global ionosphere-thermosphere model (GITM) [6], whose 10^6 states require a several-hundred-node computing cluster for real-time simulation. Real-time data assimilation based on UKF would require several million nodes, which is not feasible in the foreseeable future.

For very large scale systems, EnKF has the dubious advantage over UKF in that the number of ensemble members is not specified. However, useful guidance for the appropriate size of the EnKF ensemble based on linearized analysis is given in [7]. This analysis provides a key role in the present paper, as explained below.

This research was supported by the National Science Foundation, under grants CNS-0539053 and ATM-0325332 to the University of Michigan, Ann Arbor, USA, and by FAPEMIG, Brazil.

In Sung Kim and Dennis S. Bernstein are with Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI, USA, 48109-2140 dsbaero@umich.edu

Bruno O. S. Teixeira is with Department of Electronic Engineering, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil

In the present paper we are motivated by the need to perform data assimilation on a system such as GITM, where propagation of an ensemble throughout the data assimilation process is prohibitive. In particular, as is often the case in practice, we assume that the available measurements are time sparse, that is, occur infrequently. When measurements are available at every time step, UKF methods for systems with underlying continuous-time dynamics are given in [8]. However, these methods are prohibitive for large scale systems, and are not needed for systems in which the underlying dynamics are given in time-discretized form.

Under the assumption of time-sparse measurements, we avoid propagating the ensemble members for all of the time steps by creating an ensemble of models only when a new measurement is made available. We then propagate this ensemble into the future, thereby generating an error-covariance matrix, which, in turn, is used to create a data injection gain, which, finally, is used to assimilate the measurements at the time step at which the measurements became available. Once the measurements are assimilated, only a single simulated model is updated until new measurements become available. We call this algorithm the ensemble-on-demand Kalman filter (EnODKF). EnODKF is suboptimal since the past history of the error covariance is lost each time the ensemble is collapsed and thus disbanded. However, the computational advantages of not updating the complete ensemble throughout the process can facilitate data assimilation in applications that would otherwise be prohibitive.

The goal of the present paper is to present EnODKF and numerically investigate its properties within the context of linear systems. Nonlinear applications are readily addressed, but are deferred to future work. A key element of our investigations is the analysis of ensemble size based on the work of [7]. We provide a self-contained proof of the result of [7], which provides guidance on the size of the ensemble needed to accurately estimate the error covariance. We use this guidance within the context of EnODKF, and demonstrate the performance of EnODKF for a representative example, specifically, a heat flow problem.

II. Ensemble Kalman Filter (EnKF)

Consider the discrete-time nonlinear dynamic system

$$x_{k+1} = f(x_k, u_k, k) + w_k, \quad k = 0, 1, 2, \dots, \tag{2.1}$$

with measurements

$$y_k = h(x_k, k) + v_k, \quad k \in \mathcal{K}_{\rm d}, \tag{2.2}$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$, $y_k \in \mathbb{R}^p$, and $\mathcal{K}_{\rm d}$ denotes the set of time steps at which measurements y_k are available.

The input u_k is assumed to be known for all $k \geq 0$, and $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^p$ are uncorrelated zero-mean white noise processes with covariances Q_k and R_k, respectively. We assume that R_k is positive definite.

Equation (2.2) denotes that measurements are not available at every time step. When data are not available, the ensemble members are updated by means of a forecast step only. However, when data are available, the ensemble members are updated by both a data assimilation step and a forecast step. We now summarize the steps of the ensemble Kalman filter. For an ensemble consisting of q members at the kth step, EnKF is given by the following procedure:

For $k \notin \mathcal{K}_{\rm d}$:

Forecast Step

$$x_{k+1}^{\mathrm{f},i} = f(x_k^{\mathrm{f},i}, u_k, k) + w_k^i, \quad i = 1, \ldots, q, \tag{2.3}$$

$$x_{k+1}^{\mathrm{f}} = \frac{1}{q}\sum_{i=1}^{q} x_{k+1}^{\mathrm{f},i}. \tag{2.4}$$

For $k \in \mathcal{K}_{\rm d}$:

Data Assimilation Step

$$y_k^{\mathrm{f},i} = h(x_k^{\mathrm{f},i}, k) + v_k^i, \quad y_k^{\mathrm{f}} = \frac{1}{q}\sum_{i=1}^{q} y_k^{\mathrm{f},i}, \tag{2.5}$$

$$E_{x,k}^{\mathrm{f}} \triangleq \left[x_k^{\mathrm{f},1} - x_k^{\mathrm{f}} \quad \cdots \quad x_k^{\mathrm{f},q} - x_k^{\mathrm{f}} \right], \tag{2.6}$$

$$E_{y,k}^{\mathrm{f}} \triangleq \left[y_k^{\mathrm{f},1} - y_k^{\mathrm{f}} \quad \cdots \quad y_k^{\mathrm{f},q} - y_k^{\mathrm{f}} \right], \tag{2.7}$$

$$P_{xy,k}^{\mathrm{f}} = \frac{1}{q-1} E_{x,k}^{\mathrm{f}} (E_{y,k}^{\mathrm{f}})^{\mathrm{T}}, P_{yy,k}^{\mathrm{f}} = \frac{1}{q-1} E_{y,k}^{\mathrm{f}} (E_{y,k}^{\mathrm{f}})^{\mathrm{T}} \tag{2.8}$$

$$K_k = P_{xy,k}^{\mathrm{f}} (P_{yy,k}^{\mathrm{f}})^{-1}, \tag{2.9}$$

$$x_k^{\mathrm{da},i} = x_k^{\mathrm{f},i} + K_k \left(y_k - y_k^{\mathrm{f},i} \right), i = 1, \ldots, q, \tag{2.10}$$

$$x_k^{\mathrm{da}} = \frac{1}{q}\sum_{i=1}^{q} x_k^{\mathrm{da},i}. \tag{2.11}$$

Forecast Step

$$x_{k+1}^{\mathrm{f},i} = f(x_k^{\mathrm{da},i}, u_k, k) + w_k^i, \quad i = 1, \ldots, q, \tag{2.12}$$

$$x_{k+1}^{\mathrm{f}} = \frac{1}{q}\sum_{i=1}^{q} x_{k+1}^{\mathrm{f},i}. \tag{2.13}$$

To reproduce the process noise statistics, the noise term w_k^i, which drives the $i^{\rm th}$ ensemble member, is generated deterministically or is sampled, for instance, from a normal distribution with mean zero and covariance Q_k. Likewise, v_k^i can be sampled from a normal distribution with mean zero and covariance R_k and added to the residual $y_k - h(x_k^{\mathrm{f},i})$ in order to reproduce the measurement noise statistics.

Figure 1 illustrates EnKF. Each ensemble member is updated by time-sparse measurement data, and is propagated independently when data are not available.

III. Ensemble Size for Linear Systems

The accuracy of EnKF improves as the number of ensemble members is increased. However, a large number of ensemble members may be computationally intractable in terms of computation time and memory. Therefore, it is necessary to determine the minimum ensemble size that can adequately approximate the mean of the states and the error covariance. We now examine the required number of ensemble members for the case of linear dynamics. Specifically, we consider the linear system

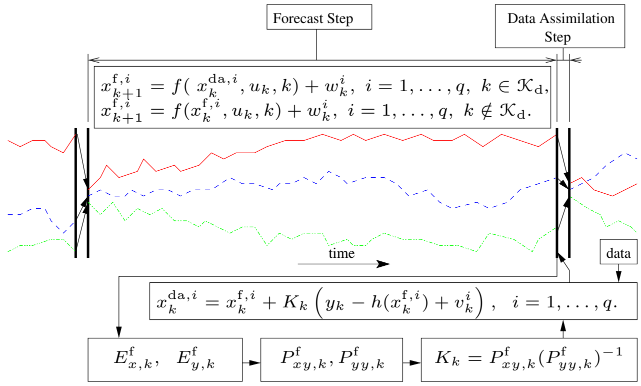

Fig. 1. Diagram of ensemble Kalman filter.

$$x_{k+1} = A_k x_k + w_k, \tag{3.1}$$

where $x_k \in \mathbb{R}^n$, $A_k \in \mathbb{R}^{n \times n}$, and $w_k \in \mathbb{R}^n$ is a random disturbance with mean zero and covariance Q_k.

Lemma III.1. *Let $\mathcal{S}_1 \subseteq \mathbb{R}^l$ and $\mathcal{S}_2 \subseteq \mathbb{R}^l$ be subspaces, and assume that*

$$\dim \mathcal{S}_1 + \dim \mathcal{S}_2 \leq l, \tag{3.2}$$

where dim *denotes dimension. Then there exists an orthogonal matrix $S \in \mathbb{R}^{l \times l}$ such that*

$$S\mathcal{S}_2 \subseteq \mathcal{S}_1^{\perp}. \tag{3.3}$$

Proof. Let $n_1 = \dim \mathcal{S}_1$ and $n_2 = \dim \mathcal{S}_2$. Let $M_1 \in \mathbb{R}^{l \times n_1}$, $M_2 \in \mathbb{R}^{l \times n_2}$ be matrices whose columns are an orthonormal basis for $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively. Next, let $M_1^{\rm c} \in \mathbb{R}^{l \times (l - n_1)}$ be a matrix composed of $l - n_1$ orthonormal vectors that are also orthogonal to each column vector of M_1, and let $M_2^{\rm c} \in \mathbb{R}^{l \times (l-n_2)}$ be defined similarly. Now define $S \triangleq [\, M_1^{\rm c} \;\; M_1] [\, M_2 \;\; M_2^{\rm c}]^{-1} \in \mathbb{R}^{l \times l}$. Hence $[\, SM_2 \;\; SM_2^{\rm c}] = [\, M_1^{\rm c} \;\; M_1]$. Since $l - n_1 \geq n_2$, it follows that $S\mathcal{S}_2 = \mathcal{R}(SM_2) \subseteq \mathcal{R}(M_1^{\rm c}) = \mathcal{S}_1^{\perp}$, where $\mathcal{R}$ denotes range. □

The following result is stated without proof in [7].

Fact III.1. *Let $x_k^{\mathrm{f},1}, \ldots, x_k^{\mathrm{f},q} \in \mathbb{R}^n$, and define*

$$H_k \triangleq [A_k x_k^{\mathrm{f},i} - A_k x_k^{\mathrm{f}}, \;\; \ldots \;\; , A_k x_k^{\mathrm{f},q} - A_k x_k^{\mathrm{f}}] \in \mathbb{R}^{n \times q}, \tag{3.4}$$

where $x_k^{\mathrm{f}} \triangleq \frac{1}{q}\sum_{i=1}^{q} x_k^{\mathrm{f},i}$. Then there exist $w_k^1, \ldots, w_k^q \in \mathbb{R}^n$ such that

$$\sum_{i=1}^{q} w_k^i = 0, \tag{3.5}$$

$$\sum_{i=1}^{q} w_k^i \left(A_k x_k^{\mathrm{f},i} - A_k x_k^{\mathrm{f}} \right)^{\mathrm{T}} = 0, \tag{3.6}$$

$$\frac{1}{q-1}\sum_{i=1}^{q} w_k^i w_k^{i\mathrm{T}} = Q_k, \tag{3.7}$$

if and only if

$$\mathrm{rank}(H_k) + \mathrm{rank}(Q_k) + 1 \leq q, \tag{3.8}$$

Now, let

$$x_{k+1}^{\mathrm{f},i} = A_k x_k^{\mathrm{f},i} + w_k^i, \quad i = 1, \ldots, q. \tag{3.9}$$

Then

$$x_{k+1}^{\mathrm{f}} = A_k x_k^{\mathrm{f}}, \tag{3.10}$$

$$P_{xx,k+1}^{\mathrm{f}} = A_k P_{xx,k}^{\mathrm{f}} A_k^{\mathrm{T}} + Q_k, \tag{3.11}$$

where $x_{k+1}^{\mathrm{f}} \triangleq \frac{1}{q}\sum_{i=1}^{q} x_{k+1}^{\mathrm{f},i}$, $P_{xx,k}^{\mathrm{f}} \triangleq \frac{1}{q-1}\sum_{i=1}^{q}\left(x_k^{\mathrm{f},i} - x_k^{\mathrm{f}}\right)\left(x_k^{\mathrm{f},i} - x_k^{\mathrm{f}}\right)^{\mathrm{T}}$, *and* $P_{xx,k+1}^{\mathrm{f}} \triangleq \frac{1}{q-1}\sum_{i=1}^{q}\left(x_{k+1}^{\mathrm{f},i} - x_{k+1}^{\mathrm{f}}\right)\left(x_{k+1}^{\mathrm{f},i} - x_{k+1}^{\mathrm{f}}\right)^{\mathrm{T}}$.

Proof. Defining

$$W_k \triangleq [w_k^1 \quad \cdots \quad w_k^q] \in \mathbb{R}^{n\times q}, \tag{3.12}$$

(3.5)-(3.7) can be written as

$$\begin{bmatrix} 1_{1\times q} \\ H_k \\ W_k \end{bmatrix} W_k^{\mathrm{T}} = \begin{bmatrix} 0 \\ 0 \\ (q-1)Q_k \end{bmatrix}, \tag{3.13}$$

where $1_{1\times q}$ is the $1\times q$ ones matrix. Letting $\mathrm{col_j}(M)$ denote the j^{th} column of M, (3.13) can be written as

$$\begin{bmatrix} 1_{1\times q} \\ H_k \\ W_k \end{bmatrix} \mathrm{col_j}(W_k^{\mathrm{T}}) = \begin{bmatrix} 0 \\ 0 \\ (q-1)\ \mathrm{col_j}(Q_k) \end{bmatrix}, \quad j = 1, \ldots, n. \tag{3.14}$$

To prove necessity, note that, using (3.14), Theorem 2.6.3 in [9] implies that, for all $j = 1, \ldots, n$,

$$\mathrm{rank}\begin{bmatrix} 1_{1\times q} \\ H_k \\ W_k \end{bmatrix} = \mathrm{rank}\begin{bmatrix} 1_{1\times q} & 0 \\ H_k & 0 \\ W_k & (q-1)\mathrm{col_j}(Q_k) \end{bmatrix} \le q. \tag{3.15}$$

Since $H_k 1_{q\times 1} = 0$, it follows that $1_{1\times q}$ is orthogonal to every row of H_k. Therefore,

$$\mathrm{rank}\begin{bmatrix} 1_{1\times q} \\ H_k \end{bmatrix} = \mathrm{rank}(H_k) + 1. \tag{3.16}$$

Furthermore, it follows from (3.14) that every row of W_k is orthogonal to every row of $[1_{1\times q}^{\mathrm{T}} \quad H_k^{\mathrm{T}}]^{\mathrm{T}}$. Finally, since $\mathrm{rank}(W_k) = \mathrm{rank}(W_k W_k^{\mathrm{T}}) = \mathrm{rank}(Q_k)$, it follows that

$$\mathrm{rank}\begin{bmatrix} 1_{1\times q} \\ H_k \\ W_k \end{bmatrix} = \mathrm{rank}(H_k) + \mathrm{rank}(Q_k) + 1 \le q. \tag{3.17}$$

To show sufficiency, let $\hat{W}_k \in \mathbb{R}^{n\times q}$ be such that $\hat{W}_k \hat{W}_k^{\mathrm{T}} = (q-1)Q_k$, and define the matrix

$$\bar{H}_k \triangleq \begin{bmatrix} 1_{1\times q} \\ H_k \end{bmatrix} \in \mathbb{R}^{(n+1)\times q}. \tag{3.18}$$

Let $\mathcal{S}_1 \triangleq \mathcal{R}(\bar{H}_k^{\mathrm{T}}) \subseteq \mathbb{R}^q$ and $\mathcal{S}_2 \triangleq \mathcal{R}(W_k^{\mathrm{T}}) \subseteq \mathbb{R}^q$. Since $\dim \mathcal{S}_1 + \dim \mathcal{S}_2 = 1 + \mathrm{rank}(H_k) + \mathrm{rank}(Q_k) \le q$, Lemma III.1 implies that there exists an orthogonal matrix $S \in \mathbb{R}^{q\times q}$ such that

$$W_k^{\mathrm{T}} = S\hat{W}_k^{\mathrm{T}}, \tag{3.19}$$

$$\bar{H}_k W_k^{\mathrm{T}} = 0, \tag{3.20}$$

$$W_k W_k^{\mathrm{T}} = \hat{W}_k S^{\mathrm{T}} S \hat{W}_k^{\mathrm{T}} = \hat{W}_k \hat{W}_k^{\mathrm{T}} = (q-1)Q_k. \tag{3.21}$$

Hence (3.13) follows. Finally, (3.10) and (3.11) follow from (3.13). □

Let

$$\mathcal{Q} \triangleq \{q : (3.8) \text{ is satisfied}\}. \tag{3.22}$$

Then, Fact III.1 shows that the minimum number of ensemble members needed to achieve (3.10), (3.11) is

$$q_{\min} \triangleq \min \mathcal{Q}. \tag{3.23}$$

Furthermore, the maximum value of q given by (3.23) is $2n+1$, where n is the number of states of the system. This value is the number of ensemble members used by UKF [5]. However, in many cases, H_k and the disturbance covariance Q_k have low rank, which means that the required ensemble size q may be substantially less than $2n+1$.

Now, we present the numerical algorithm given in [7] for generating $w_k^1, \ldots, w_k^q$ that satisfy (3.5)-(3.7); see Fact III.2. This algorithm is used for the numerical examples in the section VI. For $z_j = [z_{j,1} \ \cdots \ z_{j,j}]^{\mathrm{T}} \in \mathbb{R}^j$, define the Householder matrix $\mathcal{H}(z_j) \in \mathbb{R}^{j\times j}$ by

$$\mathcal{H}(z_j) \triangleq I - \frac{1}{1+|z_{j,j}|}\begin{bmatrix} z_{j,1} \\ \vdots \\ z_{j,j-1} \\ z_{j,j}+\mathrm{sign}(z_{j,j}) \end{bmatrix}\begin{bmatrix} z_{j,1} \\ \vdots \\ z_{j,j-1} \\ z_{j,j}+\mathrm{sign}(z_{j,j}) \end{bmatrix}^{\mathrm{T}}, \tag{3.24}$$

and let $\hat{\mathcal{H}}(z_j) \in \mathbb{R}^{j\times(j-1)}$ denote $\mathcal{H}(z_j)$ with its last column deleted. Next, define $\Gamma \in \mathbb{R}^{q\times(n+1)}$ by

$$\Gamma \triangleq \bar{H}_k^{\mathrm{T}}, \tag{3.25}$$

and let $h_k \triangleq \mathrm{rank}(H_k)$ and $q_k \triangleq \mathrm{rank}(Q_k)$.

Fact III.2. *Let $q \ge h_k + q_k + 1$ be an integer, and let $\hat{W}_k \in \mathbb{R}^{n\times q}$ satisfy $\hat{W}_k \hat{W}_k^{\mathrm{T}} = (q-1)Q_k$. Let γ be a nonzero column of Γ, and define $z_q \triangleq \gamma/\|\gamma\|_2$. Form Γ_0 by removing γ from Γ. For $i = 0, \ldots, h_k - 1$, let $\gamma_i \in \mathbb{R}^{q-i-1}$ be a nonzero column of $\hat{\mathcal{H}}^{\mathrm{T}}(z_{q-i})\Gamma_i \in \mathbb{R}^{(q-i-1)\times(n-i-1)}$, and define $z_{q-i-1} \triangleq \gamma_i/\|\gamma_i\|_2$. Remove γ_i from $\hat{\mathcal{H}}^{\mathrm{T}}(z_{q-i})\Gamma_i$ and denote the resulting matrix by Γ_{i+1}. Finally, let $\Omega \in \mathbb{R}^{(q-1-h_k)\times q_k}$ satisfy $\Omega^{\mathrm{T}}\Omega = I$. Then*

$$W_k = [w_k^1 \cdots w_k^q] \triangleq \hat{W}_k \hat{\mathcal{H}}(z_q) \cdots \hat{\mathcal{H}}(z_{q-h_k})\Omega \tag{3.26}$$

satisfies (3.5)-(3.7).

IV. Ensemble Reduction for Linear System Data Assimilation

Consider the linear system

$$x_{k+1} = A_k x_k + w_k \tag{4.1}$$

with measurements

$$y_k = C_k x_k + v_k, \quad k \in \mathcal{K}_{\mathrm{d}}, \tag{4.2}$$

where $x_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}^p$, and $\mathcal{K}_{\mathrm{d}}$ denotes the set of time steps at which measurements y_k are available. As in Section II, $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^p$ are uncorrelated zero-mean white noise processes with covariances Q_k and R_k, respectively. We assume that R_k is positive definite. For this linear system, the Kalman filter is given by the following procedure:

For $k \notin \mathcal{K}_{\mathrm{d}}$:

$$x_{k+1}^{\mathrm{f}} = A_k x_k^{\mathrm{f}}, \tag{4.3}$$

$$P_{k+1}^{\mathrm{f}} = A_k P_k^{\mathrm{f}} A_k^{\mathrm{T}} + Q_k. \tag{4.4}$$

For $k \in \mathcal{K}_{\mathrm{d}}$:

Data Assimilation Step

$$y_k^{\mathrm{f}} = C_k x_k^{\mathrm{f}}, \tag{4.5}$$

$$K_k = P_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1}, \tag{4.6}$$

$$x_k^{\mathrm{da}} = x_k^{\mathrm{f}} + K_k (y_k - y_k^{\mathrm{f}}), \tag{4.7}$$

$$P_k^{\mathrm{da}} = P_k^{\mathrm{f}} - P_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} C_k P_k^{\mathrm{f}}, \tag{4.8}$$

Forecast Step

$$x_{k+1}^{\mathrm{f}} = A_k x_k^{\mathrm{da}}, \tag{4.9}$$

$$P_{k+1}^{\mathrm{f}} = A_k P_k^{\mathrm{da}} A_k^{\mathrm{T}} + Q_k. \tag{4.10}$$

The following result shows that disturbances that do not affect the observable subspace can be ignored by the data assimilation procedure.

Proposition IV.1. *Consider the linear system*

$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} A_{1,k} & 0 \\ A_{21,k} & A_{2,k} \end{bmatrix} \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} + \begin{bmatrix} w_{1,k} \\ w_{2,k} \end{bmatrix}, \tag{4.11}$$

$$y_k = \begin{bmatrix} C_{1,k} & 0 \end{bmatrix} \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} + v_k, \tag{4.12}$$

where $A_{2,k}$ is asymptotically stable, and $R_k \triangleq \mathcal{E}(v_k v_k^{\mathrm{T}})$. Let $w_k \triangleq \begin{bmatrix} w_{1,k}^{\mathrm{T}} & w_{2,k}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$, assume w_k and v_k are uncorrelated, and define

$$Q_k \triangleq \mathcal{E}(w_k w_k^{\mathrm{T}}) = \begin{bmatrix} Q_{1,k} & Q_{12,k} \\ Q_{12,k}^{\mathrm{T}} & Q_{2,k} \end{bmatrix}, \hat{Q}_k \triangleq \begin{bmatrix} Q_{1,k} & 0 \\ 0 & 0 \end{bmatrix}. \tag{4.13}$$

Now, let $\hat{x}_{k,Q_k}$ be the state estimate of the Kalman filter that uses Q_k, and let $\hat{x}_{k,\hat{Q}_k}$ be the state estimate of the suboptimal estimator whose gain is obtained by replacing Q_k by $\hat{Q}_k$ in (4.4), (4.10). Define $\mathcal{P}_k \triangleq \mathcal{E}[(x_k - \hat{x}_{k,Q_k})(x_k - \hat{x}_{k,Q_k})^{\mathrm{T}}]$ and $\hat{\mathcal{P}}_k \triangleq \mathcal{E}[(x_k - \hat{x}_{k,\hat{Q}_k})(x_k - \hat{x}_{k,\hat{Q}_k})^{\mathrm{T}}]$, and let the corresponding costs J_{Q_k} of the Kalman filter and $J_{\hat{Q}_k}$ of the suboptimal estimator be

$$J_{Q_k} \triangleq \operatorname{tr} \mathcal{P}_k, \quad J_{\hat{Q}_k} \triangleq \operatorname{tr} \hat{\mathcal{P}}_k. \tag{4.14}$$

Assume that the Kalman filter and the suboptimal estimator have same initial conditions and initial error covariance. Then, for all k,

$$J_{Q_k} \le J_{\hat{Q}_k}. \tag{4.15}$$

Furthermore, if $Q_{12,k} = 0$, then, for all k,

$$J_{Q_k} = J_{\hat{Q}_k}. \tag{4.16}$$

Proof. We denote the gains of the Kalman filter with Q_k and the suboptimal estimator with $\hat{Q}_k$ by K_k and $\hat{K}_k$, respectively, where K_k is given by (4.6) and $\hat{K}_k$ is given by

$$\hat{K}_k = \hat{P}_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k \hat{P}_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1}, \tag{4.17}$$

$$\hat{P}_k^{\mathrm{da}} = \hat{P}_k^{\mathrm{f}} - \hat{P}_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k \hat{P}_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} C_k \hat{P}_k^{\mathrm{f}}, \tag{4.18}$$

$$\hat{P}_{k+1}^{\mathrm{f}} = A_k \hat{P}_k^{\mathrm{da}} A_k^{\mathrm{T}} + \hat{Q}_k. \tag{4.19}$$

Then the error covariance $\mathcal{P}_k$ of the Kalman filter and the pseudo-error covariance $\hat{\mathcal{P}}_k$ of the suboptimal estimator satisfy

$$\begin{aligned} \mathcal{P}_{k+1} = & A_k (I - K_k C_k) \mathcal{P}_k (I - K_k C_k)^{\mathrm{T}} A_k^{\mathrm{T}} \\ & + A_k K_k R_k K_k^{\mathrm{T}} A_k^{\mathrm{T}} + Q_k, \end{aligned} \tag{4.20}$$

$$\begin{aligned} \hat{\mathcal{P}}_{k+1} = & A_k (I - \hat{K}_k C_k) \hat{\mathcal{P}}_k (I - \hat{K}_k C_k)^{\mathrm{T}} A_k^{\mathrm{T}} \\ & + A_k \hat{K}_k R_k \hat{K}_k^{\mathrm{T}} A_k^{\mathrm{T}} + Q_k. \end{aligned} \tag{4.21}$$

Subtracting (4.20) from (4.21), adding and subtracting $A_k (I - \hat{K}_k C_k) \mathcal{P}_k (I - \hat{K}_k C_k)^{\mathrm{T}} A_k^{\mathrm{T}}$, and using $K_k = \mathcal{P}_k C_k^{\mathrm{T}} \tilde{R}_k^{-1}$, where $\tilde{R}_k \triangleq C_k \mathcal{P}_k C_k^{\mathrm{T}} + R_k$, yields

$$\begin{aligned} & \hat{\mathcal{P}}_{k+1} - \mathcal{P}_{k+1} = A_k (I - \hat{K}_k C_k)(\hat{\mathcal{P}}_k - \mathcal{P}_k)(I - \hat{K}_k C_k)^{\mathrm{T}} A_k^{\mathrm{T}} \\ & + A_k \hat{K}_k \tilde{R}_k \hat{K}_k^{\mathrm{T}} A_k^{\mathrm{T}} + A_k K_k \tilde{R}_k K_k^{\mathrm{T}} A_k^{\mathrm{T}} - A_k^{\mathrm{T}} K_k \tilde{R}_k \hat{K}_k^{\mathrm{T}} A_k^{\mathrm{T}} - A_k \hat{K}_k \tilde{R}_k K_k^{\mathrm{T}} A_k^{\mathrm{T}} \\ & = A_k (I - \hat{K}_k C_k)(\hat{\mathcal{P}}_k - \mathcal{P}_k)(I - \hat{K}_k C_k)^{\mathrm{T}} A_k^{\mathrm{T}} + A_k (\hat{K}_k - K_k) \tilde{R}_k (\hat{K}_k - K_k)^{\mathrm{T}} A_k^{\mathrm{T}} \\ & \ge 0, \end{aligned} \tag{4.22}$$

which implies (4.15).

Now, assume $Q_{12,k} = 0$. Let P_k^{f} and $\hat{P}_k^{\mathrm{f}}$ denote the forecast-step error covariance and pseudo-error covariance of the Kalman filter with Q_k and the suboptimal estimator with Q_k replaced by $\hat{Q}_k$, respectively. Next, partition P_k^{f} and $\hat{P}_k^{\mathrm{f}}$ according to (4.11) as

$$P_k^{\mathrm{f}} = \begin{bmatrix} P_{1,k}^{\mathrm{f}} & P_{12,k}^{\mathrm{f}} \\ P_{12,k}^{\mathrm{fT}} & P_{2,k}^{\mathrm{f}} \end{bmatrix}, \hat{P}_k^{\mathrm{f}} = \begin{bmatrix} \hat{P}_{1,k}^{\mathrm{f}} & \hat{P}_{12,k}^{\mathrm{f}} \\ \hat{P}_{12,k}^{\mathrm{fT}} & \hat{P}_{2,k}^{\mathrm{f}} \end{bmatrix}. \tag{4.23}$$

Similarly, define and partition P_k^{da} and $\hat{P}_k^{\mathrm{da}}$. Let P_0 denote the initial error covariance, and define the initial forecast step error covariances P_0^{f}, $\hat{P}_0^{\mathrm{f}}$ by

$$P_0^{\mathrm{f}} = \hat{P}_0^{\mathrm{f}} = P_0. \tag{4.24}$$

Now, the gains K_k and $\hat{K}_k$ are given by

$$K_k = \begin{bmatrix} P_{1,k}^{\mathrm{f}} C_{1,k}^{\mathrm{T}} \\ P_{12,k}^{\mathrm{f}} C_{1,k}^{\mathrm{T}} \end{bmatrix} V_k^{-1}, \hat{K}_k = \begin{bmatrix} \hat{P}_{1,k}^{\mathrm{f}} C_{1,k}^{\mathrm{T}} \\ \hat{P}_{12,k}^{\mathrm{f}} C_{1,k}^{\mathrm{T}} \end{bmatrix} \hat{V}_k^{-1}, \tag{4.25}$$

where $V_k \triangleq C_{1,k} P_{1,k}^{\mathrm{f}} C_{1,k}^{\mathrm{T}} + R_k$ and $\hat{V}_k \triangleq C_{1,k} \hat{P}_{1,k}^{\mathrm{f}} C_{1,k}^{\mathrm{T}} + R_k$. Using the gains K_k and $\hat{K}_k$, $P_{1,k}^{\mathrm{da}}$, $P_{12,k}^{\mathrm{daT}}$ and $\hat{P}_{1,k}^{\mathrm{da}}$, $\hat{P}_{12,k}^{\mathrm{daT}}$ are given by

$$P_{1,k}^{\mathrm{da}} = P_{1,k}^{\mathrm{f}} - P_{1,k}^{\mathrm{f}} C_{1,k}^{\mathrm{T}} V_k^{-1} C_{1,k} P_{1,k}^{\mathrm{f}}, \tag{4.26}$$

$$P_{12,k}^{\mathrm{daT}} = P_{12,k}^{\mathrm{fT}} - P_{12,k}^{\mathrm{fT}} C_{1,k}^{\mathrm{T}} V_k^{-1} C_{1,k} P_{1,k}^{\mathrm{f}}, \tag{4.27}$$

$$\hat{P}_{1,k}^{\mathrm{da}} = \hat{P}_{1,k}^{\mathrm{f}} - \hat{P}_{1,k}^{\mathrm{f}} C_{1,k}^{\mathrm{T}} V_k^{-1} C_{1,k} \hat{P}_{1,k}^{\mathrm{f}}, \tag{4.28}$$

$$\hat{P}_{12,k}^{\mathrm{daT}} = \hat{P}_{12,k}^{\mathrm{fT}} - \hat{P}_{12,k}^{\mathrm{fT}} C_{1,k}^{\mathrm{T}} V_k^{-1} C_{1,k} \hat{P}_{1,k}^{\mathrm{f}}. \tag{4.29}$$

Consequently, P_{k+1}^{f} and $\hat{P}_{k+1}^{\mathrm{f}}$ are given by

$$P_{1,k+1}^{\mathrm{f}} = A_{1,k} P_{1,k}^{\mathrm{da}} A_{1,k}^{\mathrm{T}} + Q_{1,k}, \tag{4.30}$$

$$P_{12,k+1}^{\mathrm{fT}} = A_{21,k} P_{1,k}^{\mathrm{da}} A_{1,k}^{\mathrm{T}} + A_{2,k} P_{12,k}^{\mathrm{daT}} A_{1,k}^{\mathrm{T}}, \tag{4.31}$$

$$\hat{P}_{1,k+1}^{\mathrm{f}} = A_{1,k} \hat{P}_{1,k}^{\mathrm{da}} A_{1,k}^{\mathrm{T}} + Q_{1,k}, \tag{4.32}$$

$$\hat{P}_{12,k+1}^{\mathrm{fT}} = A_{21,k} \hat{P}_{1,k}^{\mathrm{da}} A_{1,k}^{\mathrm{T}} + A_{2,k} \hat{P}_{12,k}^{\mathrm{daT}} A_{1,k}^{\mathrm{T}}. \tag{4.33}$$

Hence, $P_{1,k}^{\mathrm{f}} = \hat{P}_{1,k}^{\mathrm{f}}$, $P_{12,k}^{\mathrm{fT}} = \hat{P}_{12,k}^{\mathrm{fT}}$, and $K_k = \hat{K}_k$ for all k, which implies $\hat{x}_{k,Q_k} = \hat{x}_{k,\hat{Q}_k}$ and thus (4.16). □

Proposition IV.1 implies that, for EnKF, it is not necessary to generate $w_{2,k}^i$, which does not affect the observable subspace of (A_k, C_k).

Now, for the disturbance $\begin{bmatrix} w_{1,k}^{\mathrm{T}} & 0 \end{bmatrix}^{\mathrm{T}}$, let the corresponding matrices for H_k and Q_k be $\hat{H}_k$ and $\hat{Q}_k$. Then, assuming that $\operatorname{rank}(\hat{H}_k) \le \operatorname{rank}(H_k)$, the minimum ensemble size needed by the ensemble Kalman filter to satisfy constraints (3.5)-(3.7) can be reduced by $\operatorname{rank}(Q_k) - \operatorname{rank}(\hat{Q}_k)$.

V. Ensemble-On-Demand Kalman Filter (EnODKF)

EnKF requires in (2.3) that q ensemble members be updated in parallel at every time step whether or not data are available. When $q_{\min}$ given by (3.23) is large, real-time estimation for acceptable accuracy is computationally expensive. To partially overcome the excessive computational complexity of the ensemble Kalman filter, we consider the *ensemble on-demand Kalman filter* (EnODKF), which

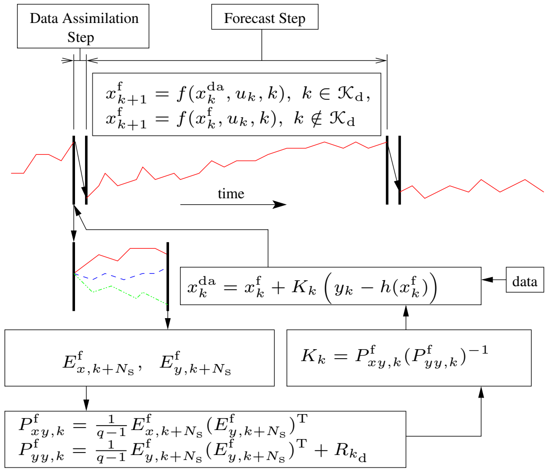

Fig. 2. Diagram of the ensemble-on-demand Kalman filter.

propagates the ensemble members over a small number $N_{\rm s}$ of steps only when data are available. EnODKF is described by following procedure: For $k \notin \mathcal{K}_{\rm d}$:

Forecast Step

$$x^{\rm f}_{k+1} = f(x^{\rm f}_k, u_k, k). \tag{5.1}$$

For $k \in \mathcal{K}_{\rm d}$:

Data Assimilation Step

$$\begin{aligned} x^{{\rm f},i}_k &= x^{\rm f}_k, \quad i = 1, \ldots, q, \\ x^{{\rm f},i}_{k+j} &= f(x^{{\rm f},i}_{k+j-1}, u_{k+j-1}, k) + w^i_{k+j-1}, \\ &\quad i = 1, \ldots, q, \quad j = 1, \ldots, N_{\rm s}, \end{aligned} \tag{5.2}$$

$$y^{{\rm f},i}_{k+N_{\rm s}} = h(x^{{\rm f},i}_{k+N_{\rm s}}, k+N_{\rm s}), \tag{5.3}$$

$$x^{\rm f}_{k+N_{\rm s}} = \frac{1}{q}\sum_{i=1}^{q} x^{{\rm f},i}_{k+N_{\rm s}}, \quad y^{\rm f}_{k+N_{\rm s}} = \frac{1}{q}\sum_{i=1}^{q} y^{{\rm f},i}_{k+N_{\rm s}}, \tag{5.4}$$

$$E^{\rm f}_{x,k+N_{\rm s}} \triangleq \left[x^{{\rm f},1}_{k+N_{\rm s}} - x^{\rm f}_{k+N_{\rm s}} \ \cdots\ x^{{\rm f},q}_{k+N_{\rm s}} - x^{\rm f}_{k+N_{\rm s}}\right], \tag{5.5}$$

$$E^{\rm f}_{y,k+N_{\rm s}} \triangleq \left[y^{{\rm f},1}_{k+N_{\rm s}} - y^{\rm f}_{k+N_{\rm s}} \ \cdots\ y^{{\rm f},q}_{k+N_{\rm s}} - y^{\rm f}_{k+N_{\rm s}}\right], \tag{5.6}$$

$$P^{\rm f}_{xy,k} = \frac{1}{q-1}E^{\rm f}_{x,k+N_{\rm s}}(E^{\rm f}_{y,k+N_{\rm s}})^{\rm T}, \tag{5.7}$$

$$P^{\rm f}_{yy,k} = \frac{1}{q-1}E^{\rm f}_{y,k+N_{\rm s}}(E^{\rm f}_{y,k+N_{\rm s}})^{\rm T} + R_{k_{\rm d}}, \tag{5.8}$$

$$K_k = P^{\rm f}_{xy,k}(P^{\rm f}_{yy,k})^{-1}, \tag{5.9}$$

$$x^{\rm da}_k = x^{\rm f}_k + K_k\left(y_k - h(x^{\rm f}_k)\right). \tag{5.10}$$

Forecast Step

$$x^{\rm f}_{k+1} = f(x^{\rm da}_k, u_k, k). \tag{5.11}$$

Figure 2 illustrates EnODKF. Each ensemble member propagates for $N_{\rm s}$ steps when data are available in order to generate an approximate error covariance. Then, the states are updated using the gain and the data at the measurement time.

Ensemble size q for EnODKF can be chosen such that $q \geq 1 + {\rm rank}(H_{k'}) + {\rm rank}(Q_{k'})$ where $k \leq k' < k + N_{\rm s}$, $k \in \mathcal{K}_{\rm d}$. Next, we should determine $N_{\rm s}$ considering the tradeoff between computation time and accuracy. That is, larger $N_{\rm s}$ ensures better accuracy as shown in Figure 5, while it requires increased computation time.

VI. Two-Dimensional Heat Conduction Example

We consider EnKF and EnODKF for the linear system (4.1) with measurements (4.2). As a baseline reference, we also compute estimates using the Kalman filter (4.3)-(4.10). For all simulations, the truth model is the model with stochastic drivers, the no data assimilation (NoDA) model is the model with the mean value of each driver, and the data assimilation (DA) model is the model with data assimilation using simulated measurements from the truth model simulation.

Consider the heat conduction in a two-dimensional plate, governed by

$$\frac{\partial T(x,y,t)}{\partial t} = \alpha\left(\frac{\partial^2 T(x,y,t)}{\partial x^2} + \frac{\partial^2 T(x,y,t)}{\partial y^2}\right) + w(x,y,t), \tag{6.1}$$

where $T(x,y,t)$ is the temperature at position (x,y) and time t, $w(x,y,t)$ represents disturbance heat sources or sinks acting on the plate, and α is the heat conduction coefficient. We discretize (6.1) over a spatial grid of size $n_x \times n_y = 20 \times 20$, where n_x and n_y denote the number of grid points in the horizontal and vertical directions, respectively. The random initial conditions $\mathcal{N}(0, I_n)$ are given, and all boundary conditions are free.

We consider two kinds of disturbance inputs, specifically, 1) n_y-independent disturbances to the left boundary edge and 2) a single disturbance to the center of the left boundary edge. Next, we consider two cases of measurements, where single measurement point is selected with different distances from the left boundary edge. We assume that measurements are available every $N_{\rm d}{=}6$ steps, and we consider $N_{\rm s}{=}1$ and $N_{\rm s}{=}4$ for the EnODKF. The disturbances and measurements are illustrated in Figure 3.

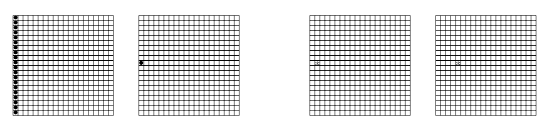

Fig. 3. Illustration of types of disturbances (left two) and measurements (right two) used for 2D heat conduction estimation. The leftmost disturbance indicates that there are 20 independent disturbances act on all of the cells of the left boundary edge.

Proposition IV.1 shows that we do not need to include in the ensemble Kalman filter the disturbance sources that do not affect the observable subspace of (A_k, C_k). Consequently, the ensemble size needed to achieve acceptable accuracy is less than $1 + {\rm rank}(H_k) + {\rm rank}(Q_k)$.

To illustrate Proposition IV.1, we consider estimation of the two-dimensional heat conduction in a square plate composed of two regions that have different heat conduction coefficients as shown in Figure 4.

We assume that the states of the $\alpha_{\rm large}$ region are observable from the measurement, and are reachable by the 10 independent disturbance sources that are in the $\alpha_{\rm large}$ region, whereas the remaining 10 disturbance sources in the $\alpha_{\rm small}$ region do not affect the $\alpha_{\rm large}$ region due to lower conductivity. Then, we perform EnKF data assimilations with the

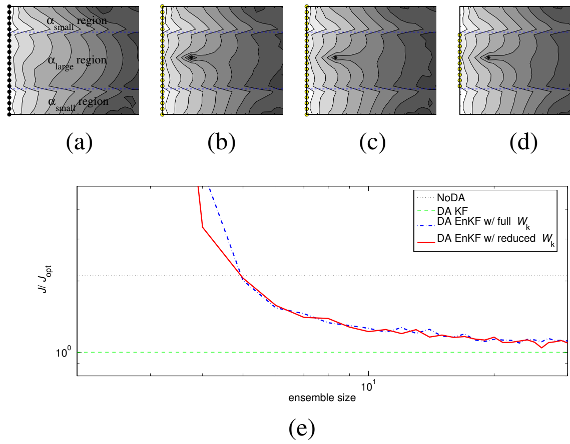

Fig. 4. Comparisons of EnKF estimators using full disturbances and reduced disturbances for 2D heat conduction on the square plate composed of regions of two different heat conduction coefficients $\alpha_{\rm small}$ and $\alpha_{\rm large}$. The $\alpha_{\rm small}$ and $\alpha_{\rm large}$ regions are shown in (a) divided by thin dashed lines, with NoDA rms error distributions. We take $\alpha_{\rm small} = 0.2\alpha_{\rm large}$. 20 filled-circles in (a) indicate the 20 independent disturbances in the truth model. (b) is the rms error distribution of the Kalman filter with the measurement at the location marked by $*$. (c) and (d) are the rms error distribution of EnKF estimators at ensemble size 30 with full 20 and 10 disturbances in the $\alpha_{\rm large}$ region, respectively. It can be seen in (e) that the errors of EnKF with full disturbance converge at around the ensemble size 20, which is far less than $1+\mathrm{rank}(H_k)+\mathrm{rank}(Q_k)$ for the entire system, which is greater than 40. Furthermore, the errors of the EnKF with fewer disturbances yields the same converged estimation accuracy as EnKF with full disturbances at around the ensemble size 20.

10 and the 20 independent disturbance sources, respectively, while increasing the number of ensemble members.

It can be seen from Figure 4(e) that the errors of EnKF with full disturbance sources converge at around an ensemble size of 20, which is less than half of $1 + \mathrm{rank}(H_k) + \mathrm{rank}(Q_k)$ for the entire system. Next, the errors of EnKF with fewer disturbance sources yields the same estimation accuracy as EnKF with full disturbance sources at an ensemble size of around 20, which means that there is no accuracy degradation with the reduced disturbance sources since the reduced disturbance sources affect the $\alpha_{\rm large}$ region while remaining 10 disturbances have minimal effect on the $\alpha_{\rm large}$ region.

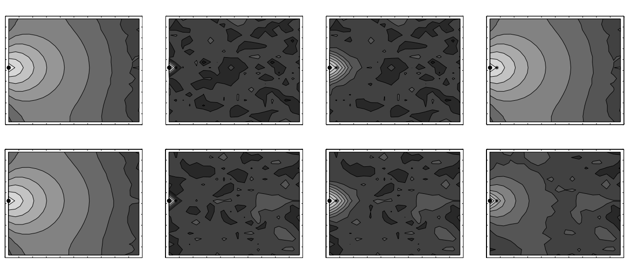

Fig. 5. 2D heat conduction estimation rms error distribution with $N_{\rm d} = 6$, $N_{\rm s} = 1$ (top) and $N_{\rm d} = 6$, $N_{\rm s} = 4$ (bottom). We assume uniform heat conductivity. 1st to 4th columns: rms error distribution of NoDA, Kalman filter, ensemble Kalman filter, and ensemble on-demand Kalman filter, respectively. There is single disturbance to the left boundary edge. Darker regions around measurement points indicate that the errors are reduced by data assimilation relative to NoDA.

Next, we compare EnKF and EnODKF. The 2D heat conduction system has slow dynamics, and the disturbances are damped out rapidly while passing through the cells. Therefore, EnODKF with $N_{\rm s} = 1$ works poorly when the measurement point and the disturbance location are different. This characteristic is illustrated in top plots of Figure 5, where all EnODKF results are identical to NoDA. However, as shown in bottom plots of Figure 5, the EnODKF with $N_{\rm s} = 4$ is effective for the cases where the measurement locations are not far from the disturbances. However, EnODKF fails to work once the measurement location is placed farther from the disturbances.

VII. Conclusion

In using EnKF, the main issues are how to perturb the system and how many ensemble members should be generated. In this paper, we showed that implementation of data assimilation by EnKF involves first identifying disturbance sources and then exciting the system using the identified disturbance sources with the ensemble size guided by the number $1 + \mathrm{rank}(H_k) + \mathrm{rank}(Q_k)$.

For large scale systems, the number $1 + \mathrm{rank}(H_k) + \mathrm{rank}(Q_k)$ may be prohibitively large for the available computing resources, and thus the reduction of computational complexity is needed. However, before enforcing the reduction of computational complexity using, for example, SVD, projection of disturbance, and, model reduction, removing unnecessary disturbance sources should be preceded in performing EnKF. We showed the effectiveness of removing unnecessary disturbances in a 2D heat conduction example with decreased computational burden and no degradation of accuracy.

Finally, we showed that EnODKF is computationally inexpensive but provides acceptable performance for systems under a single global-type disturbance.

References

[1] P. L. Houtekamer and H. L. Mitchel, “A sequential ensemble Kalman Filter for Atmospheric Data Assimilation,”, *Monthly Weather Review*, 129, pp. 796–811, 2001.

[2] L. Nerger, S. Danilov, G. Kivman, W. Hiller, and J. Schroter, “Data assimilation with the Ensemble Kalman Filter and the SEIK filter applied to a finite element model of the North Atlantic,” *Journal of Marine Systems*, 65, pp.288–298 ,2007.

[3] P. Moireau, D. Chapelle, and P. L. Tallec, “Joint state and parameter estimation for distributed mechanical systems,” *Comput. Methods Appl. Mech. Engrg.*, 197, pp.659–677, 2008.

[4] G. Evensen, “Sequential data assimilation with a nonlinear quasi-geotrophic model using Monte Carlo methods to forecast error statistics,” *Journal of Geographical Research*, 99:C5 pp. 10,143–10162, 1994.

[5] S. J. Julier and J. K. Uhlmann, “Unscented Filtering and Nonlinear Estimation,” *Proc. of the IEEE*, 92(3):pp. 401–422, March 2004.

[6] A. J. Ridley, Y. Deng, and G. Tóth. “The global ionosphere thermosphere model,” *Journal of Atmospheric and Solar-Terrestrial Physics*, 68:pp. 839–864, 2006.

[7] D. T. Pham, “Stochastic Methods for Sequential Data Assimilation in Strongly Nonlinear Systems,” *Monthly Weather Rev.*, 129: pp.1194–1207, 2000.

[8] S. Särkkä, “On Unscented Kalman Filtering for State Estimation of Continuous-Time Nonlinear Systems,” *IEEE Trans. on Auto. Contr.*, 52:9 pp.1631–1641, 2007.

[9] D. S. Bernstein, *Matrix Mathematics*, Princeton University Press, 2005.