NAVAL POSTGRADUATE SCHOOL

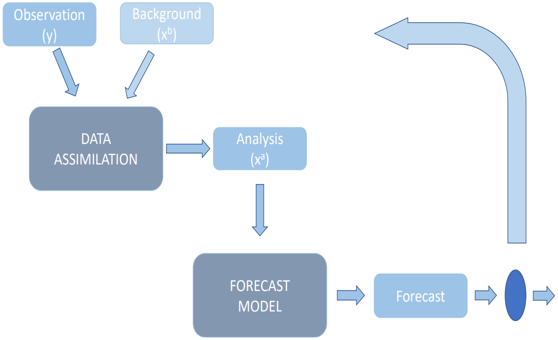
Data Assimilation for Models with a Sparse Error Covariance

Wei Kang
U.S. Naval Postgraduate School

Liang Xu
U.S. Naval Research Laboratory

2018 Workshop on Sensitivity Analysis and Data Assimilation in Meteorology and Oceanography

Numerical Weather Prediction



Data Assimilation

System Model

$$
\begin{aligned}
x_{k+1} &= \mathcal{M}(x_k) + \eta_k, \quad x_k \in \mathbb{R}^n, \quad \eta_k \sim \text{model uncertainty}, Q \\
y_k &= \mathcal{H}(x_k) + \delta_k, \quad y_k \in \mathbb{R}^m, \quad \delta_k \sim \text{sensor noise}, R
\end{aligned}
$$

Linearization

$$
\begin{aligned}
x_{k+1} &= M_k x_k + \cdots \\
y_k &= H_k x_k + \cdots
\end{aligned}
$$

Both n and m are very large. In daily operations, only a small part of sensor data is used.

Challenges

4D-Var

- It is an effective method to provide estimation results with an affordable computational load.
- The method does not provide information about error covariance.
- It requires tangent linear models and adjoint models.

EnKF

- EnKF does not require tangent linear model and adjoint model
- It contains partial information about error statistics
- Undersampling and rank deficiency
- Filter divergence, inbreeding, spurious correlations

Sparsity-based Filters

Sparsity-based filters: The goal is to avoid rank deficiency, provide more error covariance information, and achieve granularity control for optimal parallelism.

A variety of parallel computing architectures are available; and new technologies are being developed rapidly.

- Multi-core CPU
- General-purpose GPU
- Clusters or massively parallel computing
- Grid computing
- Application-specific integrated curcuits
- ......

Sparsity of error covariance

Ensemble	Sparse Covariance
X X X X X	X
X X X X X	X X
X X X X X	X X X
X X X X X	X X X X
X X X X X	0 X X X X
X X X X X	0 0 X X X X
X X X X X	0 0 0 X X X X
X X X X X	0 0 0 0 X X X X
X X X X X	0 0 0 0 0 X X X X
X X X X X	0 0 0 0 0 0 X X X X
X X X X X	0 0 0 0 0 0 0 X X X X
X X X X X	0 0 0 0 0 0 0 0 X X X X
X X X X X	0 0 0 0 0 0 0 0 0 X X X X
X X X X X	0 0 0 0 0 0 0 0 0 0 X X X X
X X X X X	0 0 0 0 0 0 0 0 0 0 0 X X X X
X X X X X	0 0 0 0 0 0 0 0 0 0 0 0 X X X X
X X X X X	0 0 0 0 0 0 0 0 0 0 0 0 0 X X X X
X X X X X	0 0 0 0 0 0 0 0 0 0 0 0 0 0 X X X X
X X X X X	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 X X X X
X X X X X	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 X X X X

Sparsity based methods

- Approximately sparse error covariance

$$N_{sp} = \text{maximum number of nonzero entries in columns}$$
$$\mathcal{I}_i(P) = \text{indices of nonzero entries in the } i\text{th-column}$$

- Component-based numerical model

$$\mathcal{M}(x_k^{sp}; \mathcal{I}) \text{ or } \mathcal{M}^{comp}$$
$$\mathcal{I} = \text{ indices of entries to be evaluated}$$

A progressive approach

Assume

$$M_k P_k M_k^T = P_k + \Delta P_{k+1}$$

To estimate ΔP_{k+1}, assume

$$M_{k+1} = I + \Delta M_k$$

$$x_{k+1} = \mathcal{M}(x_k) = x_k + \Delta(x_k)$$

Then

$$\begin{aligned} M_k P_k M_k^T &= (I + \Delta M_k) P_k (I + \Delta M_k^T) \\ &= P_k + \Delta M_k P_k + (\Delta M_k P_k)^T + \cdots \\ &\approx \left(\mathcal{M}(x_k + \delta P_k) - \mathcal{M}(x_k)\right)/\delta \\ &+ \left(\mathcal{M}(x_k + \delta P_k) - \mathcal{M}(x_k)\right)^T/\delta - P_k \end{aligned}$$

Progressive KF

Background $x_{k|k}$ and $P^{sp}_{k|k}$ (sparse covariance approximation)

Forecast

$$\begin{aligned}
x_{k+1|k} &= \mathcal{M}(x_{k|k}) \\
y_{k+1|k} &= \mathcal{H}(x_{k|k}) \\
P^{sp}_{k+1|k} &= \left(\mathcal{M}^{comp}(x^{sp}_{k|k} + \delta P^{sp}_{k|k}) - \mathcal{M}^{comp}(x^{sp}_{k|k})\right)/\delta \\
&\quad + \left(\mathcal{M}^{comp}(x^{sp}_{k|k} + \delta P^{sp}_{k|k}) - \mathcal{M}^{comp}(x^{sp}_{k|k})\right)^T/\delta \\
&\quad - P^{sp}_{k|k} + Q
\end{aligned}$$

Analysis

$$\begin{aligned}
K &= P^{sp}_{k+1|k} H^T_{k+1} (H_{k+1} P^{sp}_{k+1|k} H^T_{k+1} + R)^{-1} \\
P^{sp}_{k+1|k+1} &= (I - K H_{k+1}) P^{sp}_{k+1|k} \\
x_{k+1|k+1} &= x_{k+1|k} + K(y_{k+1} - y_{k+1|k})
\end{aligned}$$

Computational load

Progressive KF Full model	number of model components evaluation
$\mathcal{M}(x_k)$ $\mathcal{M}\left(x_k + \delta P_k(:, i)\right)$ $i = 1, 2, \cdots, n$	$(n+1)nN_p$ N_p- progressive steps
Progressive KF Component-based model	
$\mathcal{M}(x_k)$ $\mathcal{M}\left(x_k + \delta P_k(:, i), \mathcal{I}_i(P)\right)$ $i = 1, 2, \cdots, n$	$(N_{sp}+1)nN_p$
Ensemble KF	
$\mathcal{M}(x_k^i)$ $i = 1, 2, \cdots, N_{ens}$	$N_{ens}n$

Avoid rank deficiency and achieve granularity control

Ensemble

```
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
X X X X X
```

Rank: N_{ens}

Size: n x N_{ens}

Sparse Covariance

```
X
XX
XXXX
XXXXX
0000XX
00000XX
000000XX
0000000XX
00000000XX
000000XXXXX
0000000XXXXX
00000000XXXXX
00000000000XXX
000000000000XXX
0000000000000XXX
00000000000000XXX
000000000000000XXX
0000000000000000XXX
00000000000000000XXX
000000000000000000XXX
```

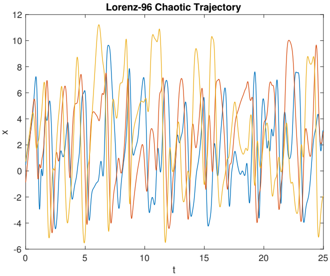
- **P has full rank**
- **N_{sp} is a variable**
- **Tasks can be grouped in different size:**
  Coarse-grained, medium-grained, fine-grained
- **P localization is straightforward**

Lorenz-96 model

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F, \quad i = 1, 2, \cdots, m$$
$$x_{m+1} = x_1$$

Discretization - 4th-order RK

$$x_k = \mathcal{M}(x_{k-1})$$
$$\Delta t = 0.025$$
$$F = 8$$
$$m = 40$$



A comparison

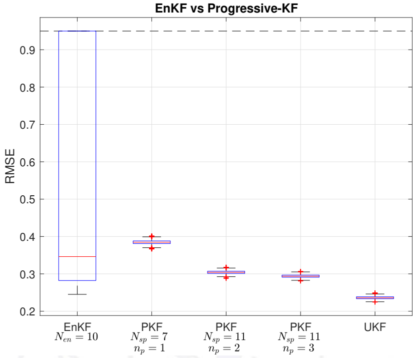
$N = 1000$ initial states in $[-1\ 1]$ - uniform distribution.

$N_{filter} = 4000$ filter steps

$m = 20$ measurement locations

$R = I$

Filter	Size	CMPT EVAL
EnKF	$N_{ens} = 10$	400
P-KF	$N_{sp} = 7$, $N_p = 1$	320
P-KF	$N_{sp} = 11$, $N_p = 2$	480x2
P-KF	$N_{sp} = 11$, $N_p = 3$	480x3



A comparison

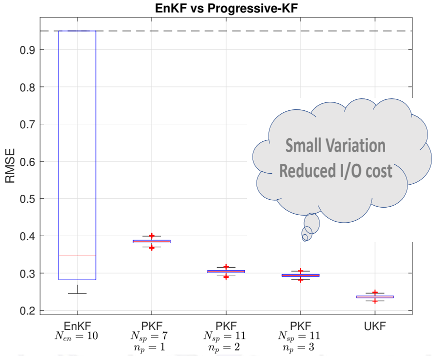
$N = 1000$ initial states in $[-1\ 1]$ - uniform distribution.

$N_{filter} = 4000$ filter steps

$m = 20$ measurement locations

$R = I$

Filter	Size	CMPT EVAL
EnKF	$N_{ens} = 10$	400
P-KF	$N_{sp} = 7$, $N_p = 1$	320
P-KF	$N_{sp} = 11$, $N_p = 2$	480x2
P-KF	$N_{sp} = 11$, $N_p = 3$	480x3



Unscented KF (UKF)

σ-points $x^i_{k|k}$, $0 \leq i \leq 2n$

$$x^0_{k|k} = x_{k|k}$$

Forecast $x^i_{k+1|k} = \mathcal{M}(x^i_{k|k}), \quad y^i_{k+1|k} = \mathcal{H}(x^i_{k|k}), \quad 0 \leq i \leq 2n$

$$\bar{x}_{k+1} = \sum_{i=0}^{2n} w_i x^i_{k+1|k}, \quad \bar{y}_{k+1} = \sum_{i=0}^{2n} w_i y^i_{k+1|k}$$

$$P_{k+1|k} = \sum_{i=0}^{2n} w_i \Delta x^i_{k+1} (\Delta x^i_{k+1})^T + Q$$

$$\Delta x^i_{k+1} = x^i_{k+1|k} - \bar{x}_{k+1}$$

$$w_0 = \frac{\kappa}{n+\kappa}, \; w_i = \frac{\kappa}{2(n+\kappa)}$$

UKF (Cont.)

Analysis

$$P^{yy} = \sum_{i=0}^{2n} w_i \Delta y^i_{k+1} (\Delta y^i_{k+1})^T + R, \quad \Delta y_{k+1} = y^i_{k+1|k} - \bar{y}_{k+1}$$

$$P^{xy} = \sum_{i=0}^{2n} w_i \Delta x^i_{k+1} (\Delta y^i_{k+1})^T$$

$$K P^{yy} = P^{xy}$$

$$x_{k+1|k+1} = \bar{x}_{k+1} + K(y_{k+1} - \bar{y}_{k+1})$$

Update

$$P_{k+1|k+1} = P_{k+1|k} - K(P^{xy})^T$$

$$x^i_{k+1|k+1} = x_{k+1|k+1} + \sqrt{(n+\kappa) P_{k+1|k+1}}, \quad i = 1, 2, \cdots, n$$

$$x^i_{k+1|k+1} = x_{k+1|k+1} - \sqrt{(n+\kappa) P_{k+1|k+1}}, \quad i = n+1, \cdots, 2n$$

Sparsity of square root matrix

Theorem (S. Toledo). If P is a symmetric positive definite matrix. The amount of storage for a Cholesky decomposition of P is $O(n + 2\eta(P))$, where $\eta(P)$ is the number of nonzero entries in P.

Assumption: The sparsity patterns of P and $(\sqrt{P})^f$ are known.

Sparse UKF

Sparse σ-points

$$x^0_{k|k} = x_{k|k}$$

$$\sigma^i,\ \mathcal{I}_i \text{ (sparsity index)} \qquad 1 \le i \le n$$

Forecast

$$x^0_{k+1|k} = \mathcal{M}(x^0_{k|k}),$$

$$x^i_{k+1|k} = \mathcal{M}^{comp}(x^0_{k|k} + \sigma^i), \qquad x^{i+n}_{k+1|k} = \mathcal{M}^{comp}(x^0_{k|k} - \sigma^i)$$

$$y^i_{k+1|k} = \mathcal{H}(x^i_{k+1|k} \triangleright_{\mathcal{I}_i} x^0_{k+1|k}), \qquad 1 \le i \le 2n$$

$$\bar{x}_{k+1} = \sum_{i=0}^{2n} w_i (x^i_{k+1|k} \triangleright_{\mathcal{I}_i} x^0_{k+1|k}), \quad \bar{y}_{k+1} = \sum_{i=0}^{2n} w_i y^i_{k+1|k}$$

$$P^{sp}_{k+1|k} = \sum_{i=0}^{2n} w_i \left(\Delta x^i_{k+1} (\Delta x^i_{k+1})^T \right)^{sp} + Q$$

$$w_0 = \frac{\kappa}{n+\kappa},\ w_i = \frac{\kappa}{2(n+\kappa)},\ \Delta x^i_{k+1} = x^i_{k+1|k} \triangleright_{\mathcal{I}_i} x^0_{k+1|k} - \bar{x}_{k+1}$$

$x^{sp}_1 \triangleright_{\mathcal{I}} x_2$ - merging operation.

Sparse UKF (Cont.)

Analysis

$$P^{yy} = \sum_{i=0}^{2n} w_i \Delta y^i_{k+1} (\Delta y^i_{k+1})^T, \quad \Delta y^i_{k+1} = y^i_{k+1|k} - \bar{y}_{k+1}$$

$$P^{xy} = \sum_{i=0}^{2n} w_i \Delta x^i_{k+1} (\Delta y^i_{k+1})^T$$

$$K P^{yy} = P^{xy}$$

$$x_{k+1|k+1} = \bar{x}_{k+1} + K(y_{k+1} - \bar{y}_{k+1})$$

Update

$$P^{sp}_{k+1|k+1} = P^{sp}_{k+1|k} - \left(K(P^{xy})^T\right)^{sp}$$

$$\sigma^i, \mathcal{I}_i \sim \sqrt{(n+\kappa) P^{sp}_{k+1|k+1}}, \qquad i = 1, 2, \cdots, n$$

A comparison

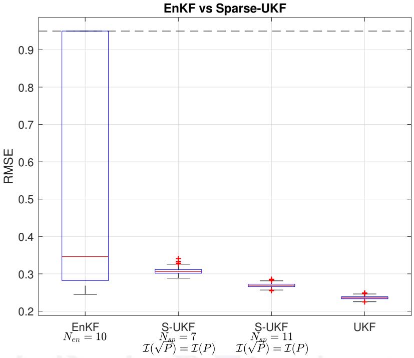
$N = 1000$ initial states in $[-1\ 1]$ - uniform distribution.

$N_{filter} = 4000$ filter steps

$m = 20$ measurement locations

$R = I$

Filter	Size	CMPT EVAL
EnKF	$N_{ens} = 10$	400
S-UKF	$N_{sp} = 7$	600
S-UKF	$N_{sp} = 11$	920



Sparsity-based Filters

A comparison

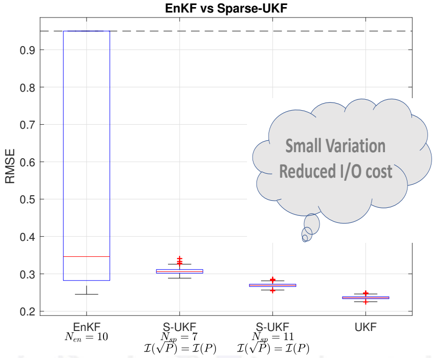
$N = 1000$ initial states in $[-1\ 1]$ - uniform distribution.

$N_{filter} = 4000$ filter steps

$m = 20$ measurement locations

$R = I$

Filter	Size	CMPT EVAL
EnKF	$N_{ens} = 10$	400
S-UKF	$N_{sp} = 7$	600
S-UKF	$N_{sp} = 11$	920



Summary

	Progressive KF	Sparse UKF
Full rank covariance	✓	✓
Small variation	✓	✓
Reduced I/O cost	✓	✓
$P_{k+1} \approx P_k + \Delta P_k$	✓	
Cholesky decomposition		✓

Summary

	Progressive KF	Sparse UKF
Full rank covariance	✓	✓
Small variation	✓	✓
Reduced I/O cost	✓	✓
$P_{k+1} \approx P_k + \Delta P_k$	✓	
Cholesky decomposition		✓

Thank you!