Kalman Filtering in Aeronautics and Astronautics

An Overview of Technology and Challenges

With special thanks to J. Chandrasekhar, H. Palanthandalam, and M. Santillo

What is a Kalman Filter ?

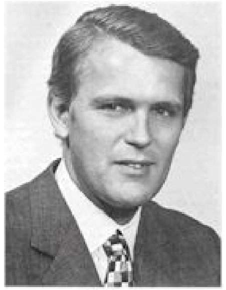
- Merging models with data to obtain estimates of states
  - Data assimilation

- Stochastically optimal observer

- Model-based filter

Development of the Kalman Filter

- **Seminal Paper**
  - **R. E. Kalman, *"A New Approach to Filtering and Prediction Problems*," Journal of Basic Engineering, Vol. 85D, pp. 35—45, 1960.**



Born 1930, Budapest, Hungary

- **R. Kalman, J. Guid. Contr. Dynamics, 2003:**
  - ***"the discovery of the Kalman filter came about through a single, gigantic, persistent mathematical exercise*."**

- **Key points:**
  1. **"I simply defined a stochastic signal source consisting of a linear system and discrete white noise"**
  2. **"I … establish[ed] for myself first the obvious relations and then the precise equivalence between transfer functions and linear vector differential equations."**
  3. **"No one imagined that the end result would be that simple."**
     - **Recursive filter**

- **Kalman and Bucy, 1961: Continuous-time case**

Kalman Filtering Applications

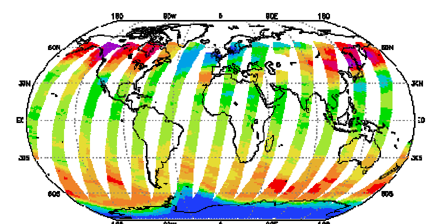
- Aerospace Applications

Guidance

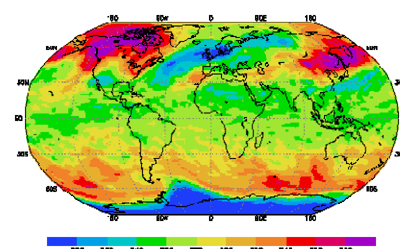
Tracking

  - Patriot missile, GPS, spacecraft

- Other Applications
  - Biological systems
    - Drug concentrations
  - Stock trading
  - Terrestrial weather
  - Space weather

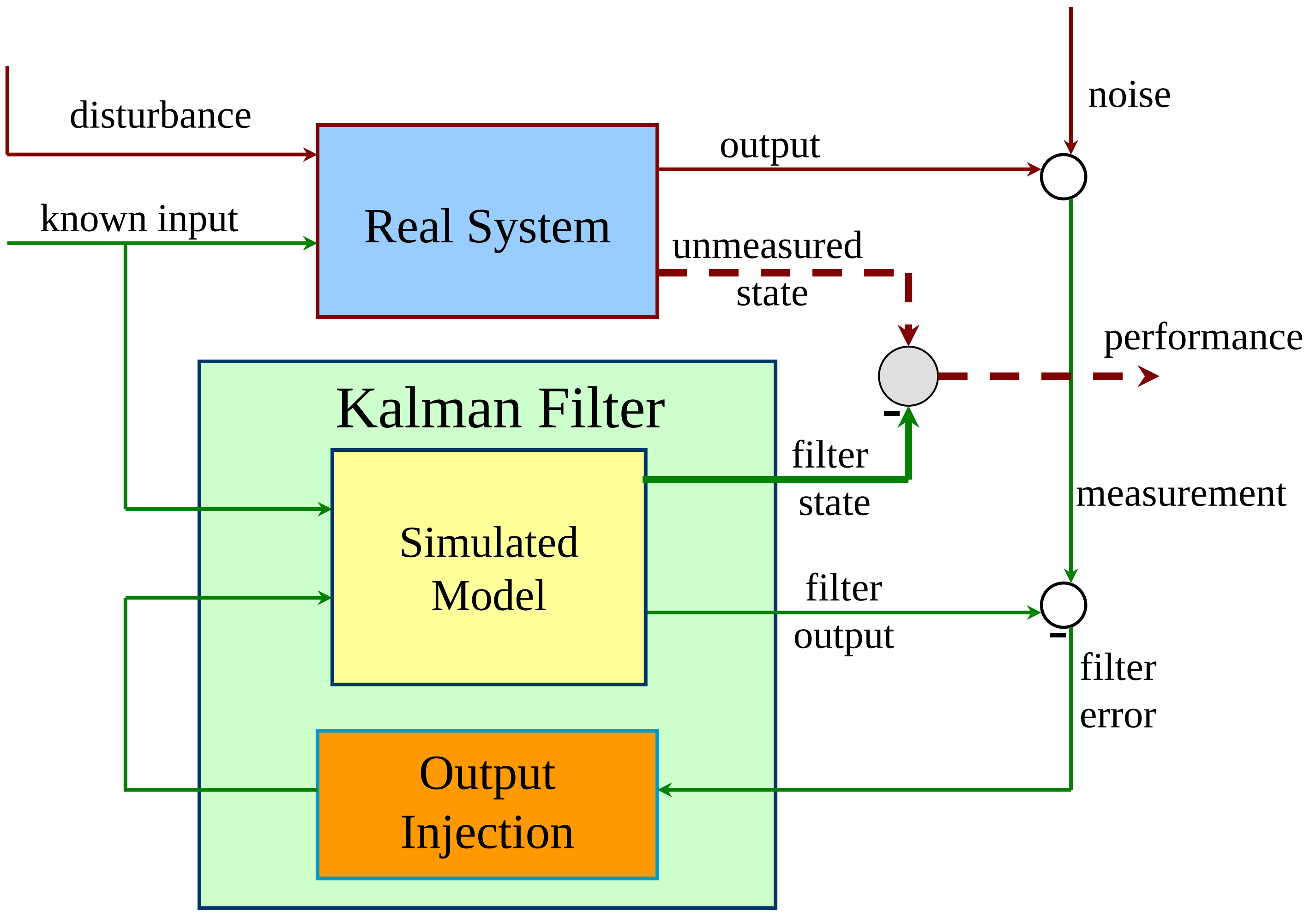


Ozone concentration data (measurements)



Global field after data assimilation

Data Assimilation



Linear Stochastic System

- Time-varying linear system

$$
\begin{aligned}
x_{k+1} &= A_k x_k + B_k u_k + H_k d_k + w_k \\
y_k &= C_k x_k + v_k
\end{aligned}
$$

$x_k \in \mathbb{R}^{n_k}$	State variable	Unknown	Stochastic
$y_k \in \mathbb{R}^{l_k}$	Measurement	Known	Stochastic
$w_k \in \mathbb{R}^{n_{k+1}}$	Unmodeled driver	Unknown	Stochastic
$v_k \in \mathbb{R}^{l_k}$	Sensor noise	Unknown	Stochastic
$u_k \in \mathbb{R}^{p_k}$	Modeled driver	Known	Deterministic
$d_k \in \mathbb{R}^{r_k}$	Unmodeled driver	Unknown	Deterministic

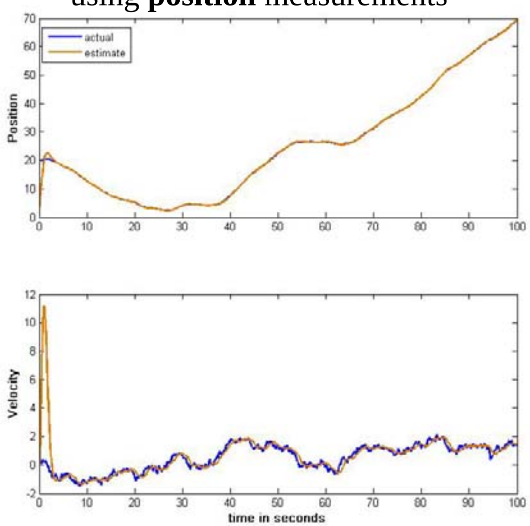
- The dimension of the state can be time varying
  - A_k is not necessarily square !!

Data Assimilation

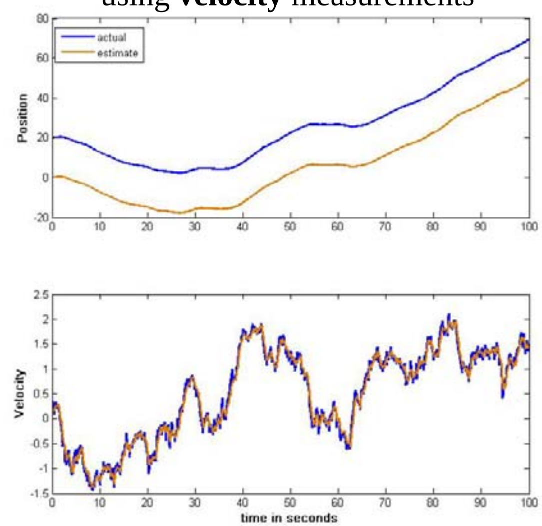
- Filtering
  - Use measurements $y_0, y_1, \cdots, y_k$ to determine the optimal estimate $\hat{x}_k$ of x_k

- Rigid body

Position and velocity estimates using **position** measurements



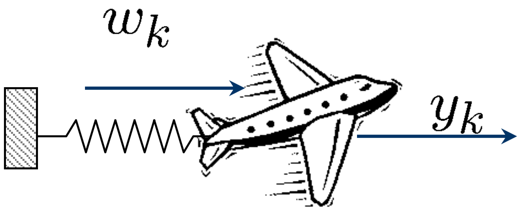
Position and velocity estimates using **velocity** measurements



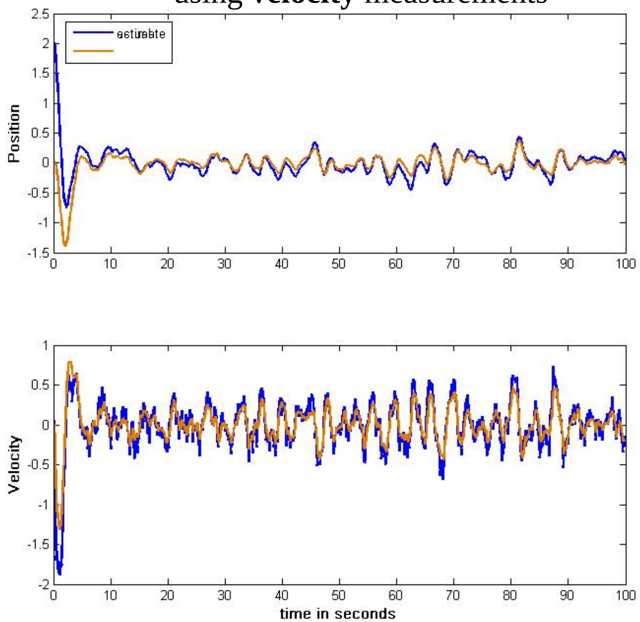
Position is not observable through velocity

Filtering

– Mass-Spring-Damper (MKC) system



Estimates of position and velocity using **velocity** measurements



Position is observable through velocity

Estimates of position improve due to observability

Optimal Estimator

- $\hat{x}_k$ = estimate of x_k
- Cost function: $J_k = \mathcal{E}[(L_k e_{k+1})^\top (L_k e_{k+1})] = \text{tr}(P_{k+1} M_k)$

$$P_k \triangleq \mathcal{E}[e_k e_k^\top], \quad M_k \triangleq L_k^\top L_k$$

  – $e_k \triangleq x_k - \hat{x}_k$
  – $\mathcal{E}[\cdot]$ is the expected value operator

- Obtain optimal estimate $\hat{x}_{k+1}$ of x_{k+1} that minimizes J_k using measurements $Y_k = \{y_0, \ldots, y_k\}$

- The optimal minimum variance estimate of x_{k+1} is

$$\hat{x}_{k+1} = \mathcal{E}[x_{k+1} | Y_k]$$

  – Optimal for arbitrary dynamics and statistics

Optimal Estimator

- Assumptions
  - Deterministic drivers are known ($d_k = 0$)
  - w_k and v_k are zero mean white Gaussian processes with covariances Q_k and R_k, respectively
  - Initial state x_0 is Gaussian with known mean $\bar{x}_0$ and known variance $\text{var}(x_0) = \mathcal{E}[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^\top]$
  - w_k and v_k are uncorrelated (*for convenience*)
- Guarantees
  - Innovation $\tilde{y}_k \triangleq y_k - \widehat{y}_k$
  - $\widehat{y}_k$ = Estimated measurement
  - $\tilde{y}_0, \ldots, \tilde{y}_k$ are mutually independent (white)
  - x_{k+1} and $\tilde{y}_k$ are jointly Gaussian

Kalman Filter

- Due to the independence of the innovation sequence

$$\mathcal{E}[x_{k+1}|Y_k] = \mathcal{E}[x_{k+1}|\tilde{Y}_k] = \mathcal{E}[x_{k+1}|\tilde{y}_k] + \mathcal{E}[x_{k+1}|\tilde{Y}_{k-1}] - E[x_{k+1}]$$

- Since x_{k+1} and $\tilde{y}_k$ are jointly Gaussian

$$\mathcal{E}[x_{k+1}|\tilde{y}_k] = \mathcal{E}[x_{k+1}] + \mathcal{E}[x_{k+1}\tilde{y}_k^\top](\mathcal{E}[\tilde{y}_k\tilde{y}_k^\top])^{-1}\tilde{y}_k$$

- Using linear dynamics

$$\mathcal{E}[x_{k+1}|\tilde{Y}_{k-1}] = \mathcal{E}[A_k x_k + B_k u_k + w_k|\tilde{Y}_{k-1}] = A_k\hat{x}_k + B_k u_k$$

$$\mathcal{E}[x_{k+1}\tilde{y}_k^\top] = A_k P_k C_k^\top$$

$$\mathcal{E}[\tilde{y}_k\tilde{y}_k^\top] = C_k P_k C_k^\top + R_k$$

Kalman Filter

- Estimator dynamics (One-step estimator)

$$\hat{x}_{k+1} = A_k \hat{x}_k + B_k u_k + K_k \underbrace{(y_k - \hat{y}_k)}_{\text{data assimilation}}$$

$$\hat{x}_0 = \bar{x}_0$$

$$K_k = A_k P_k C_k^\top (C_k P_k C_k^\top + R_k)^{-1}$$

- The optimal filter gain depends on the error covariance P_k
- The error covariance is propagated by

$$\underbrace{P_{k+1}}_{\text{Uncertainty measure}} = \underbrace{A_k P_k A_k^\top + Q_k}_{\text{Open-loop dynamics}} - \underbrace{A_k P_k C_k^\top (C_k P_k C_k^\top + R_k)^{-1} C_k P_k A_k^\top}_{\text{Uncertainty reduction due to filtering}}$$

  - Set $P_0 = \mathcal{E}[e_0 e_0^\top] = \mathrm{var}(x_0) = \mathcal{E}[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^\top]$
  - Riccati difference equation

Data Assimilation

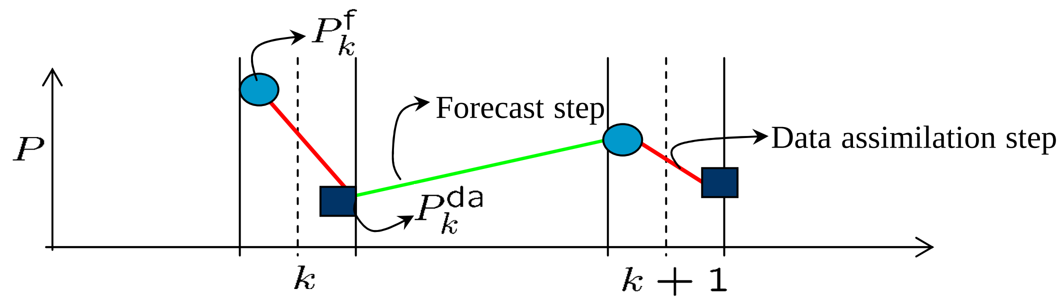
- Two-step optimal estimator
  - Equivalent to the one-step filter

$$
\begin{aligned}
x^{\mathsf{f}}_{k+1} &= A_k x^{\mathsf{da}}_k + B_k u_k \\
x^{\mathsf{da}}_k &= x^{\mathsf{f}}_k + K_k (y_k - y^{\mathsf{f}}_k) \\
y^{\mathsf{f}}_k &= C_k x^{\mathsf{f}}_k \\
K_k &= P^{\mathsf{f}}_k C^{\mathsf{T}}_k (C_k P^{\mathsf{f}}_k C^{\mathsf{T}}_k + R_k)^{-1} \\
P^{\mathsf{da}}_k &= P^{\mathsf{f}}_k - P^{\mathsf{f}}_k C^{\mathsf{T}}_k (C_k P^{\mathsf{f}}_k C^{\mathsf{T}}_k + R_k)^{-1} C_k P^{\mathsf{f}}_k \\
P^{\mathsf{f}}_{k+1} &= A_k P^{\mathsf{da}}_k A^{\mathsf{T}}_k + Q_k
\end{aligned}
$$

- Forecast (physics) update
- Data assimilation update
- Data assimilation covariance update
- Forecast covariance update

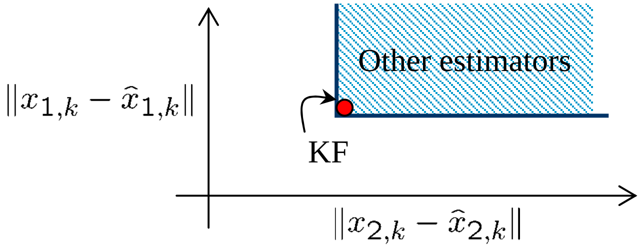
data assimilation error covariance: $P^{\mathsf{da}}_k \triangleq \mathcal{E}[e^{\mathsf{da}}_k (e^{\mathsf{da}}_k)^{\mathsf{T}}]$ — data assimilation error

forecast error covariance: $P^{\mathsf{f}}_k \triangleq \mathcal{E}[e^{\mathsf{f}}_k (e^{\mathsf{f}}_k)^{\mathsf{T}}]$ — forecast error



Kalman Filter Properties

- Optimal estimate $\widehat{x}_k$ of x_k
  - Does not depend on the error weighting L_k
    - Kalman filter provides optimal estimates of all states



Globally Pareto optimal

- Recursive update of the filter
  - At every step only the most recent measurement is used

- The optimal estimate $\widehat{x}_k$ of x_k is unbiased
  - $\mathcal{E}[x_k - \widehat{x}_k] = 0$

Kalman Filter Properties

- Under white Gaussian $w_k,\ v_k$ and x_0, the optimal estimator for a linear system is linear
- The filter gain K_k and error covariance P_k
  - Do not depend on $x_k,\ \hat{x}_k,\ y_k,\ w_k$ and v_k
  - Depend only on $A_k,\ C_k,\ Q_k,\ R_k$
- The filter gains $K_0, \ldots, K_k$ can be determined offline
- Next: Enforce a linear structure but relax assumptions on $w_k,\ v_k$ and x_0

Fixed-Structure Estimator

- Assumptions
  - $w_k, \; v_k$ and x_0 can be non-Gaussian
  - w_k and v_k are uncorrelated (*for convenience*)
- Objective : Obtain the linear minimum variance estimate of x_k

Linear Fixed-Structure Estimators

- One-step estimator

$$\begin{aligned} \hat{x}_{k+1} &= A_k \hat{x}_k + B_k u_k + K_k (y_k - \hat{y}_k) \\ \hat{y}_k &= C_k \hat{x}_k \\ \text{(cost function)} \quad J_k &= \mathcal{E}[(L_k e_{k+1})^{\mathsf{T}} (L_k e_{k+1})] = \mathrm{tr}(P_{k+1} M_k) \\ P_k &\triangleq \mathcal{E}[e_k e_k^{\mathsf{T}}] \end{aligned}$$

- Two-step estimator

$$\begin{aligned} x^{\mathrm{f}}_{k+1} &= A_k x^{\mathrm{da}}_k + B_k u_k \\ x^{\mathrm{da}}_k &= x^{\mathrm{f}}_k + K_k (y_k - y^{\mathrm{f}}_k) \\ y^{\mathrm{f}}_k &= C_k x^{\mathrm{f}}_k \\ \text{(cost function)} \quad J_k &= \mathcal{E}[(L_k e^{\mathrm{da}}_k)^{\mathsf{T}} (L_k e^{\mathrm{da}}_k)] = \mathrm{tr}(P^{\mathrm{da}}_k M_k) \end{aligned}$$

- Determine K_k to minimize J_k

Linear Estimator Error Covariances

- One-step estimator

$$J_k = \mathrm{tr}(P_{k+1} M_k)$$

$$P_{k+1} = (A_k - K_k C_k) P_k (A_k - K_k C_k)^\top + Q_k + K_k R_k K_k^\top$$

- Two-step estimator

$$J_k = \mathrm{tr}(P_k^{\mathrm{da}} M_k)$$

$$P_k^{\mathrm{da}} = (I - K_k C_k) P_k^{\mathrm{f}} (I - K_k C_k)^\top + K_k R_k K_k^\top$$

$$P_{k+1}^{\mathrm{f}} = A_k P^{\mathrm{da}} A_k^\top + Q_k$$

- Set $\dfrac{\partial J_k}{\partial K_k} = 0$ to obtain the optimal filter gain K_k

Optimal Linear Estimator

- One-step optimal linear estimator

$$P_{k+1} = A_k P_k A_k^\top - A_k P_k C_k^\top (C_k P_k C_k^\top + R_k)^{-1} C_k P_k A_k^\top + Q_k$$

$$K_k = A_k P_k C_k^\top \left(C_k P_k C_k^\top + R_k \right)^{-1}$$

- Two-step optimal linear estimator

$$P_k^{\mathsf{da}} = (I - K_k C_k) P_k^{\mathsf{f}}$$

$$P_{k+1}^{\mathsf{f}} = A_k P_k^{\mathsf{da}} A_k^\top + Q_k$$

$$K_k = P_k^{\mathsf{f}} C_k^\top \left(C_k P_k^{\mathsf{f}} C_k^\top + R_k \right)^{-1}$$

- The one-step and two-step linear minimum variance filters are equivalent

- Provides optimal linear minimum variance estimates for non-Gaussian $w_k,\ v_k$ and x_0

LTI Case

- $n_k = n,\ A_k = A,\ C_k = C,\ Q_k = Q,\ R_k = R$

- Kalman filter

$$P_{k+1} = AP_kA^\top - AP_kC^\top(CP_kC^\top + R)^{-1}CP_kA^\top + Q$$

$$K_k = AP_kC^\top\left(CP_kC^\top + R\right)^{-1}$$

  – The optimal filter gain K_k is time varying

- If (A, Q) is stabilizable and (A, C) is detectable then

$$\lim_{k\to\infty} P_k = P,\ \lim_{k\to\infty} K_k = K$$

- P satisfies the discrete algebraic Riccati equation

$$P = APA^\top - APC^\top(CPC^\top + R)^{-1}CPA^\top + Q$$

$$K = APC^\top(CPC^\top + R)^{-1}$$

LTI Case

- Steady-state error dynamics

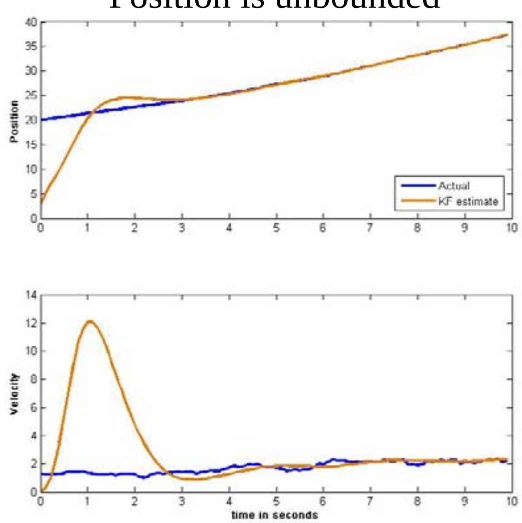
$$e_{k+1} = (A - KC)e_k + w_k + K_k v_k$$

- $A - KC$ is asymptotically stable
- Stochastically optimal LTI observer

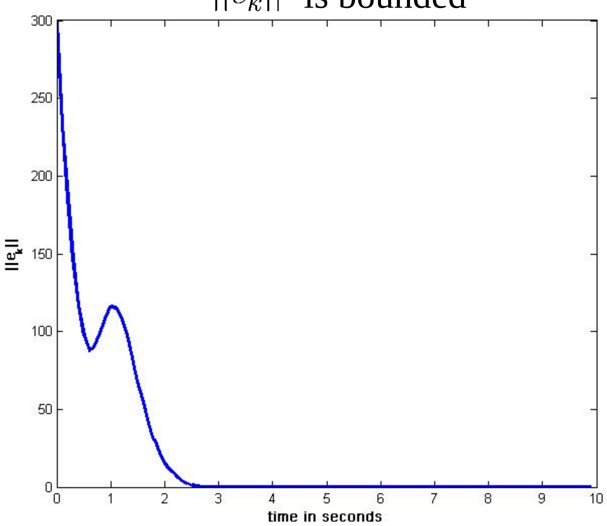
Rigid body example

Estimate position and velocity using **position** measurements

Position is unbounded



$||e_k||$ is bounded



- Error e_k is bounded even if x_k is not bounded

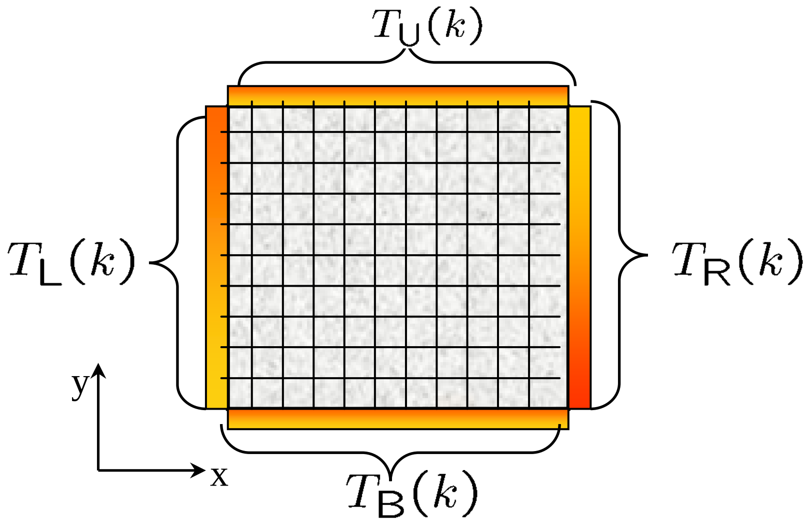
2D- Heat Conduction Example

- Equation of motion

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

$T = T(x, y, t)$ = Temperature

α = Thermal diffusivity



- Discretize PDE using finite-volume method

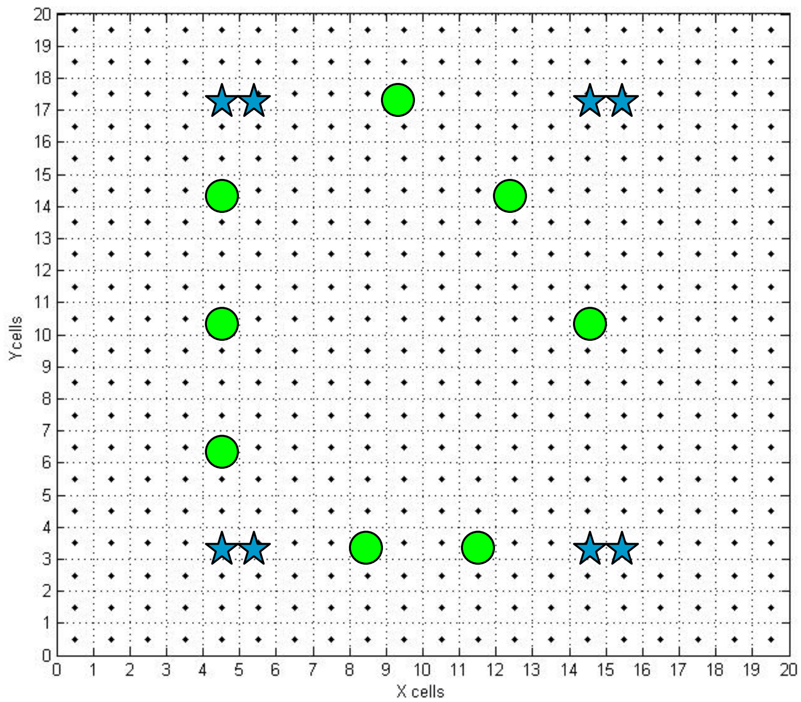
$$x_k = [T_{1,1}(k)\ T_{1,2}(k)\ \cdots\ T_{n,m}(k)]^T$$

$$x_{k+1} = Ax_k + Bu_k$$

Boundary conditions $u_k = [T_L(k)\ T_R(k)\ T_U(k)\ T_B(k)]^T$

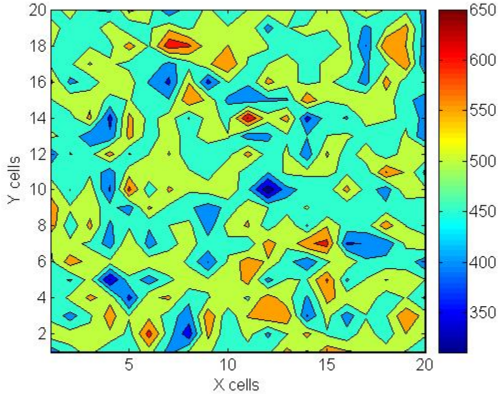
Truth Model

- Initial temperature distribution is distributed randomly with mean 500 K
- Unknown heat sources/sinks are placed at points indicated by ★
- Temperature measurements are obtained at points indicated by ●

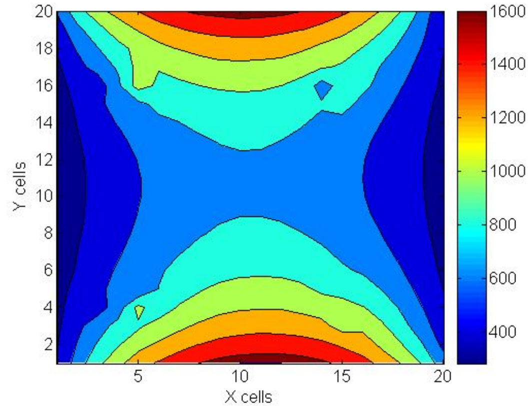


- Grid size = 20 x 20
- Dimension of state vector x_k = 400
- Boundary conditions are known
  - Sinusoidal
  - Uniform over each side

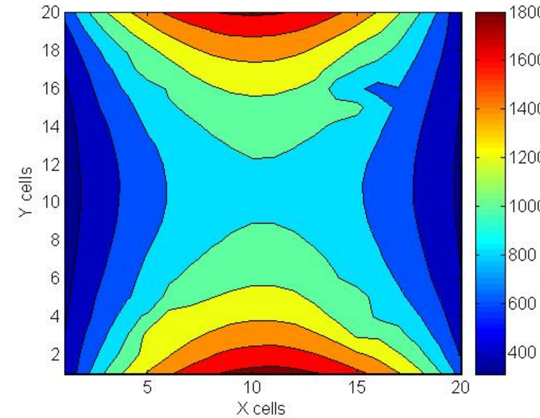
Truth Model – Temperature



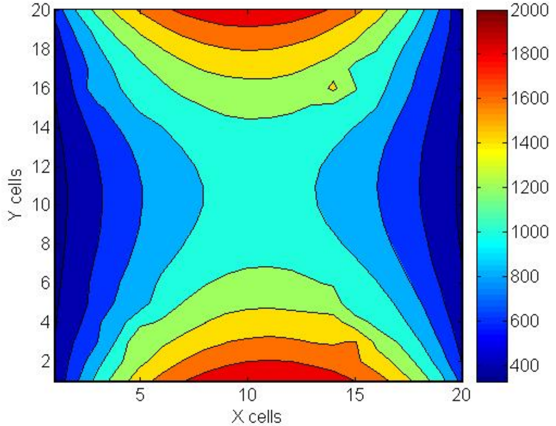
t=0 s



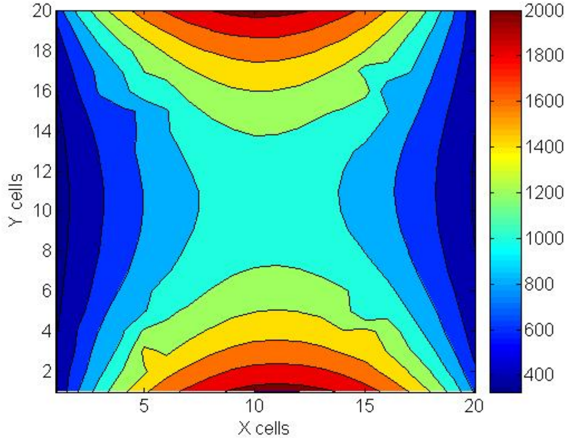
t=25 s



t=50 s



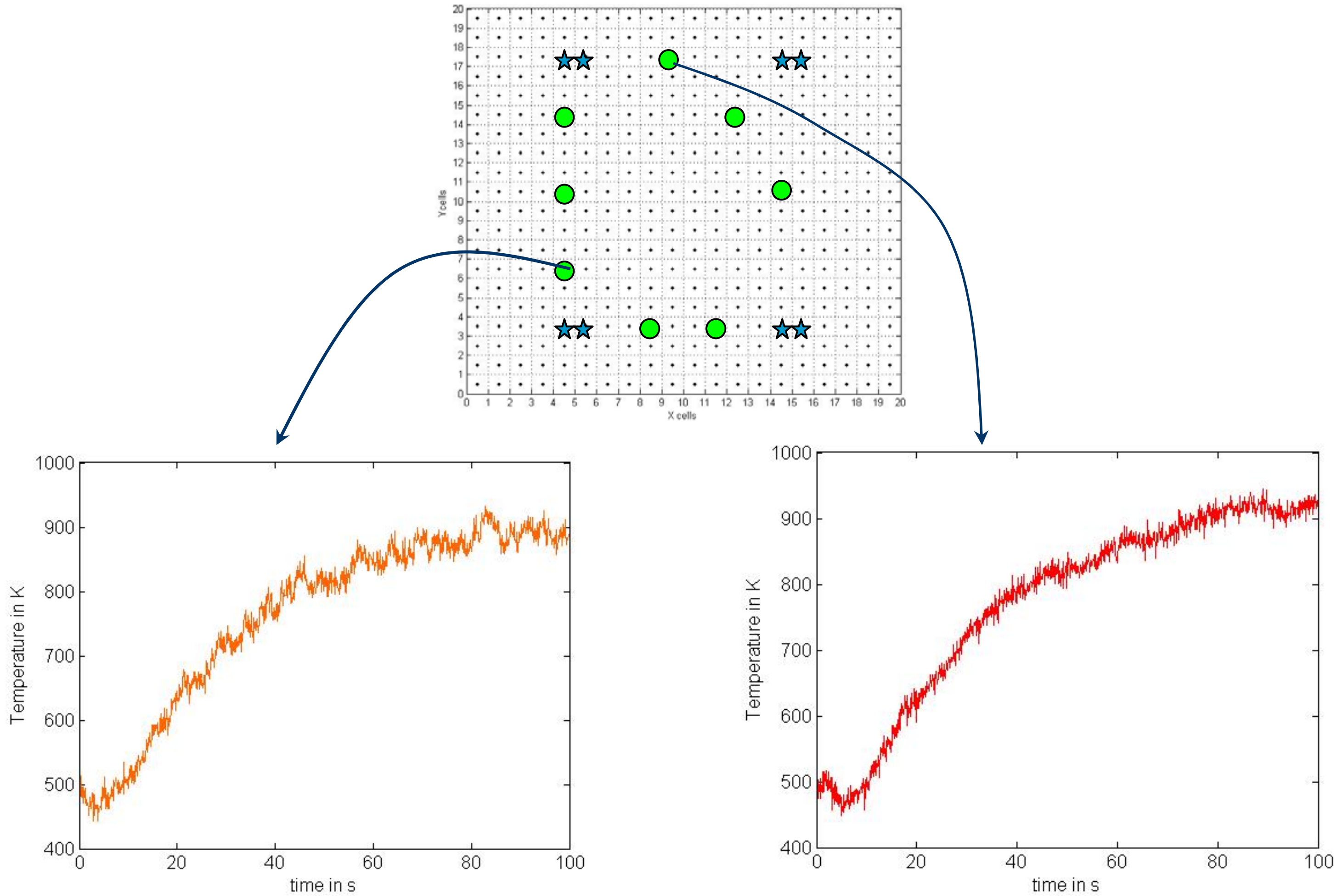
t=75 s



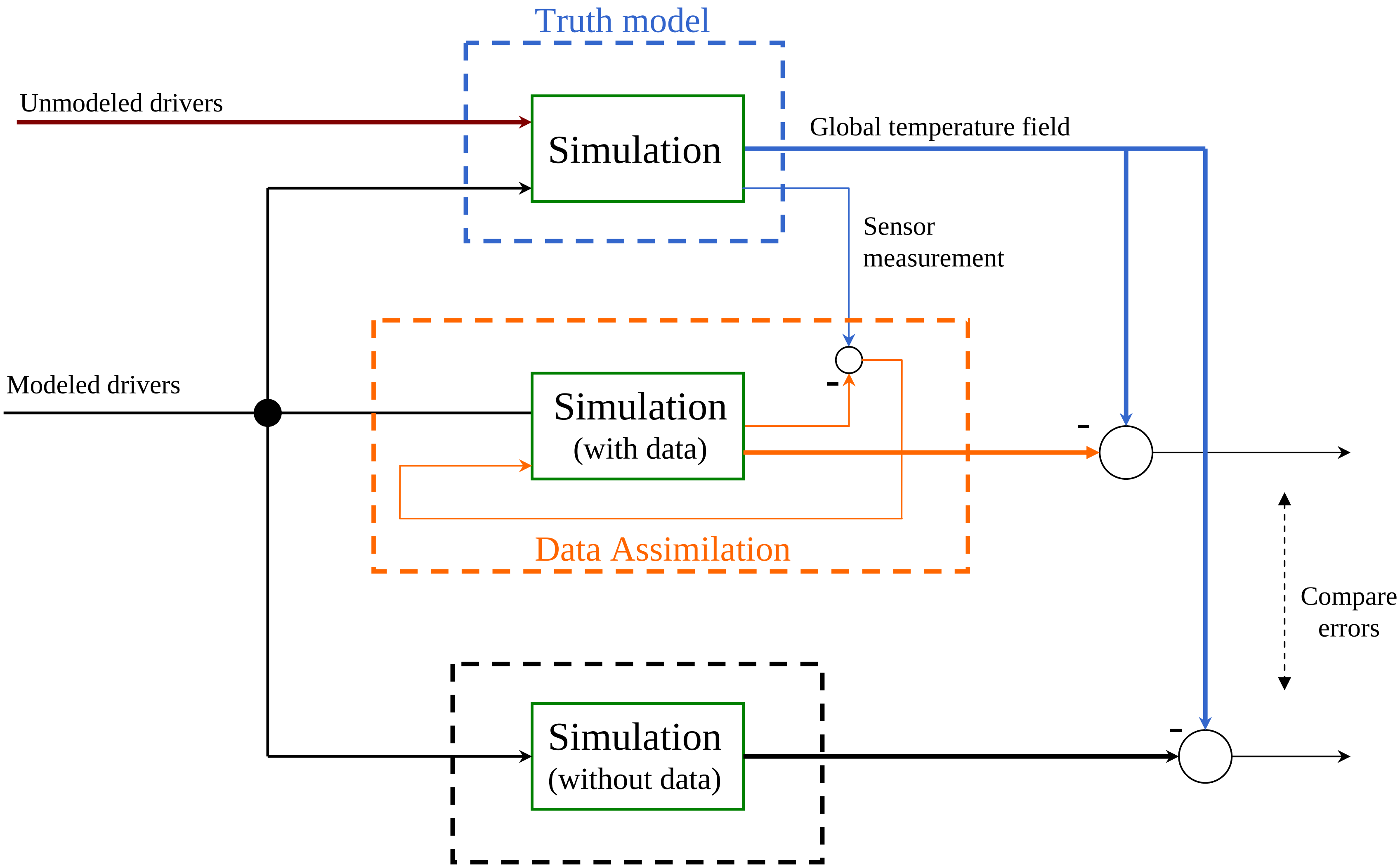
t=100 s

Truth Model Measurements

- Temperature measured by sensors

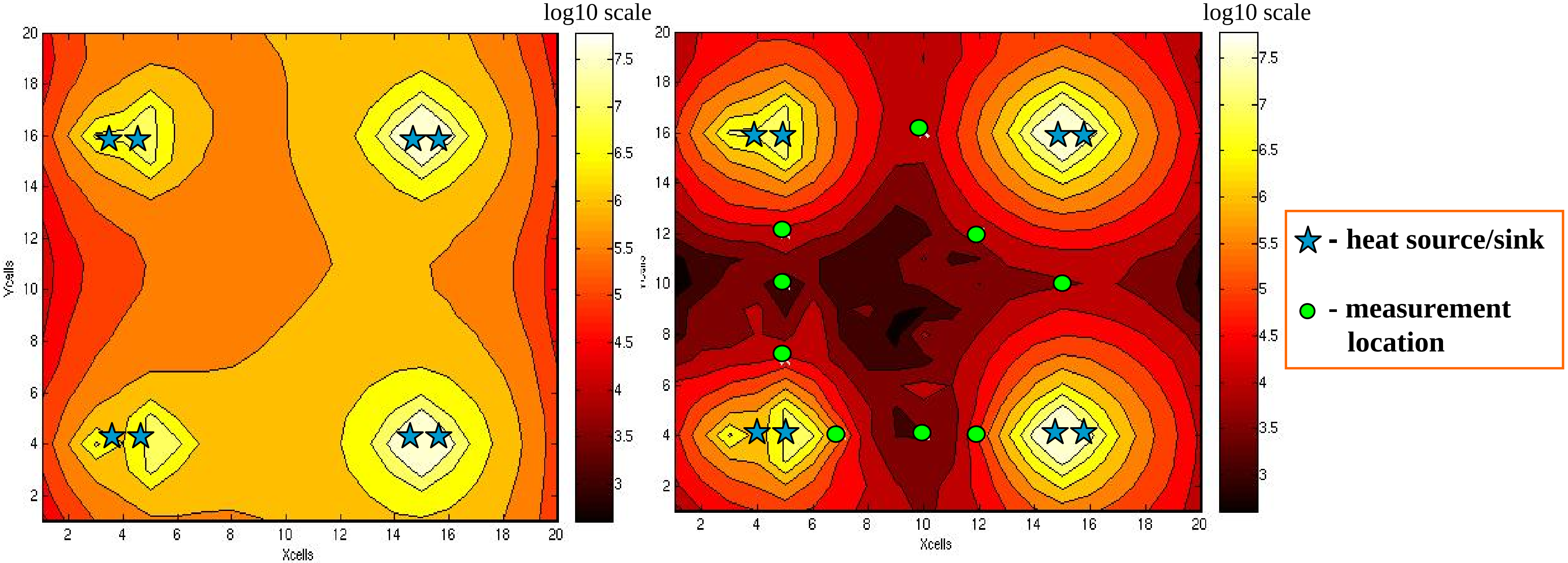


Data Assimilation Performance



Data Assimilation Results

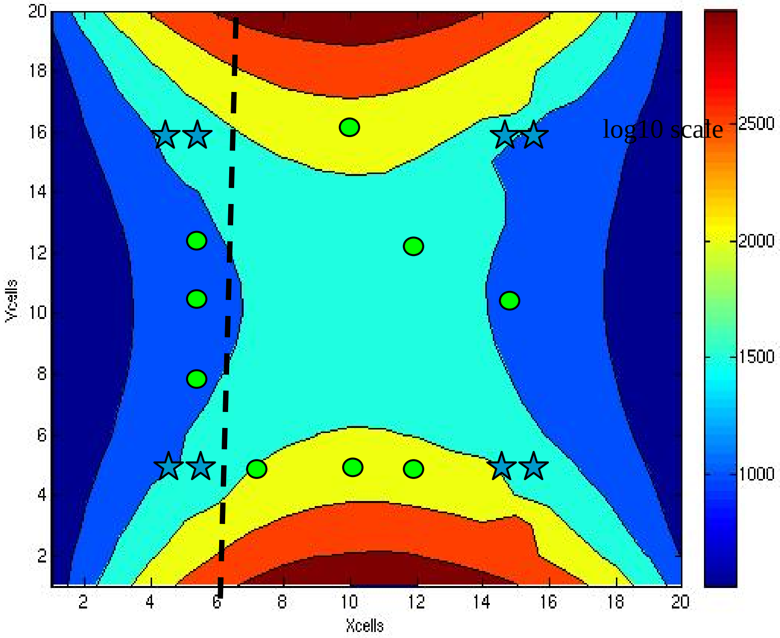
- Temperature distribution is not steady due to unmodeled drivers and time-varying boundary condition
- Error covariance and filter gain reach steady state
- Sum of the squares of error in temperature estimates between truth model (modeled drivers = known boundary conditions)



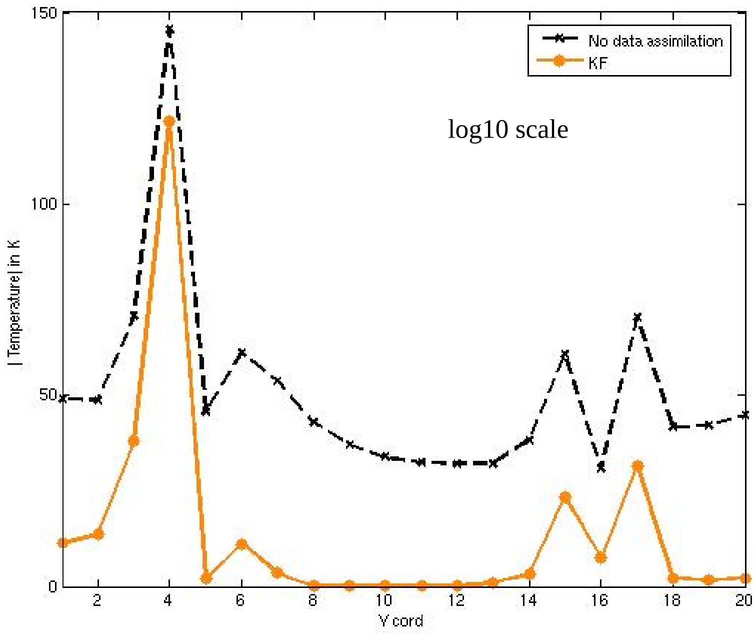
|| Truth – Simulation without data || **|| Truth – Data assimilation ||**

Data Assimilation Results

- Compare temperature profile determined by the KF with the truth model across cross sections



Temperature distribution of the plate at t=50 s



Absolute value of the error in temperature profile along X=6 at t=50 s

Extensions of the Basic Kalman Filter

1) Local data injection

2) w_k and v_k correlation

3) Unknown statistics of random variables w_k, v_k and x_0(skip)

4) Unknown deterministic input d_k

5) Known nonlinear dynamics

6) Unknown A_k and C_k

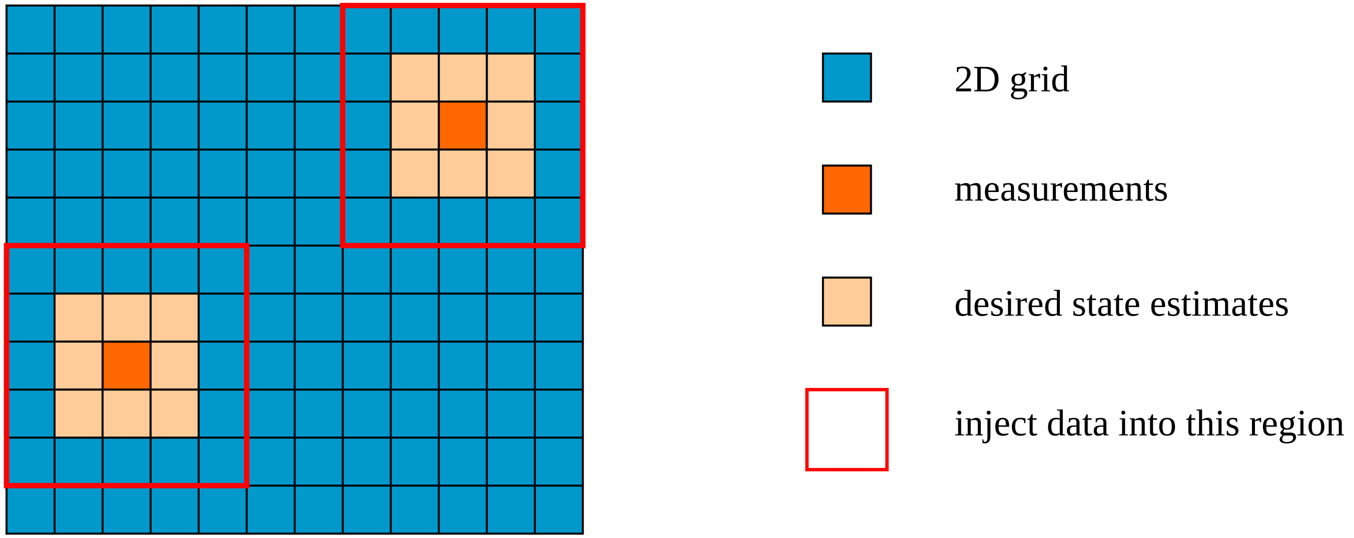
7) High-dimensional systems

8) Physical Constraints (skip)

Local Data Injection

Optimal Partial State Estimation

- Motivation
  - Kalman filter uses full data injection
  - Data might be effective in a subregion only
  - Updating all the states in a parallel multi-processor architecture is difficult



KF with Spatially Local Output Injection

- Estimate only states in the range of L_k
  - Minimize $J_k = \mathcal{E}[(L_k e_k)^\top (L_k e_k)]$

- Update specific state estimates
  - One-step estimator

$$\hat{x}_{k+1} = A_k \hat{x}_k + B_k u_k + \Gamma_k K_k (y_k - \hat{y}_k), \Gamma = \begin{bmatrix} 0 \\ I \\ 0 \end{bmatrix}$$

  - Two-step estimator

$$x^{\mathrm{f}}_{k+1} = A_k x^{\mathrm{da}}_k + B_k u_k \quad \longrightarrow \text{Forecast update}$$

$$x^{\mathrm{da}}_k = x^{\mathrm{f}}_k + \Gamma_k K_k (y_k - y^{\mathrm{f}}_k) \quad \longrightarrow \text{Data assimilation update}$$

  - Inject data from all measurements into state estimates in the range of Γ_k (Γ_k has full rank)

- The one-step and two-step optimal estimators are not equivalent

Optimal Linear Estimator

- Propagate a modified error covariance
  - One-step case

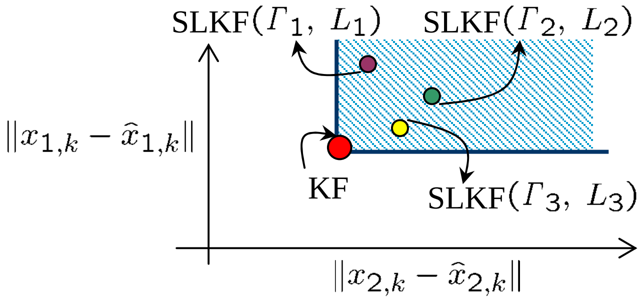
additional term

$$
\begin{aligned}
P_{k+1} &= A_k P_k A_k^\top - A_k P_k C_k^\top \hat{R}_k^{-1} C_k P_k A_k^\top + Q_k - \pi_{k\perp}^\top A_k P_k C_k^\top \hat{R}_k^{-1} C_k P_k A_k^\top \pi_{k\perp} \\
K_k &= (\Gamma_k^\top M_k \Gamma_k) \Gamma_k^\top M_k A_k P_k C_k \hat{R}_k^{-1} \\
\pi_{k\perp} &= I_n - \pi_k, \quad \pi_k = \Gamma_k (\Gamma_k^\top \Gamma_k)^{-1} \Gamma_k^\top, \quad \hat{R}_k = C_k P_k C_k^\top + R_k
\end{aligned}
$$

  - Two-step case

$$
\begin{aligned}
P_{k+1}^{\mathsf{f}} &= A_k P_k^{\mathsf{da}} A_k^\top + Q_k \\
K_k &= (\Gamma_k^\top M_k \Gamma_k) \Gamma_k^\top M_k P_k C_k \left(C_k P_k C_k^\top + R_k \right)^{-1} \\
P_k^{\mathsf{da}} &= (I - \Gamma_k K_k C_k) P_k^{\mathsf{f}} (I - \Gamma_k K_k C_k)^\top + \Gamma_k K_k R_k K_k^\top \Gamma_k^\top
\end{aligned}
$$

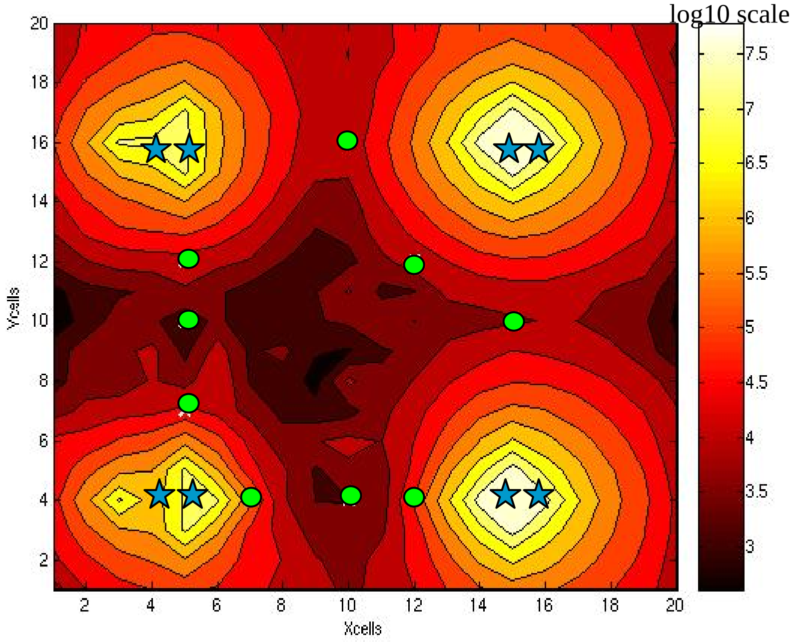
- The optimal estimates depend on the error weighting L_k



SLKF is not globally Pareto optimal

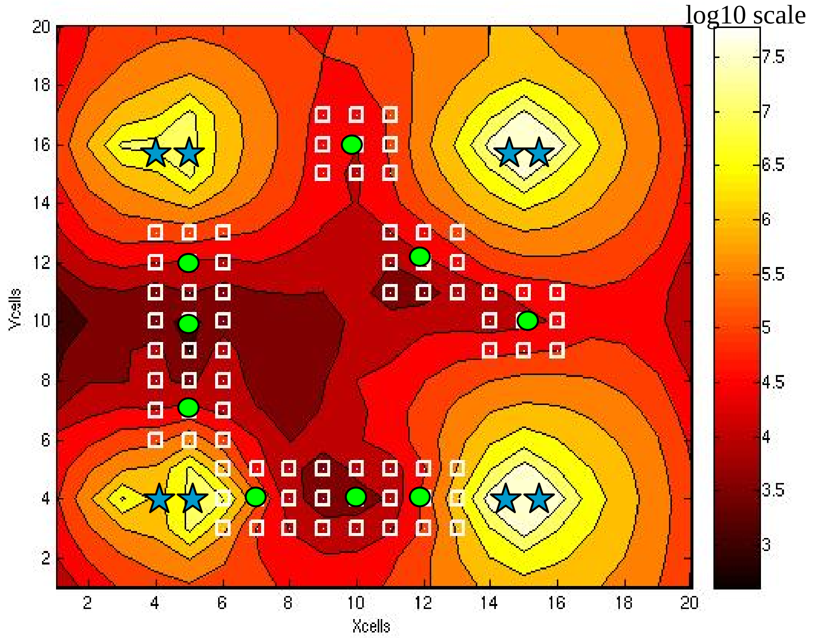
2D-Heat Conduction Example

- Compare SLKF and KF performance
  - 2D Heat Conduction Example



Kalman filter

Data is injected into all the cells

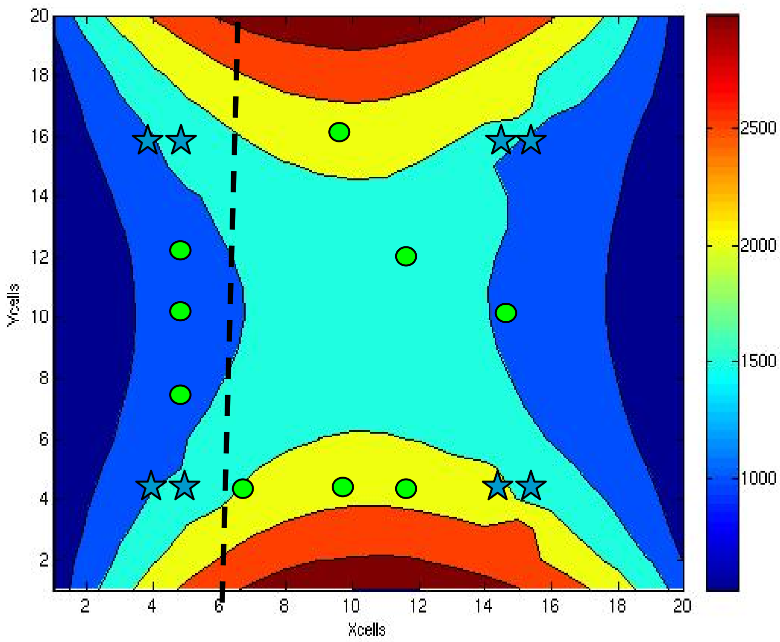


SLKF

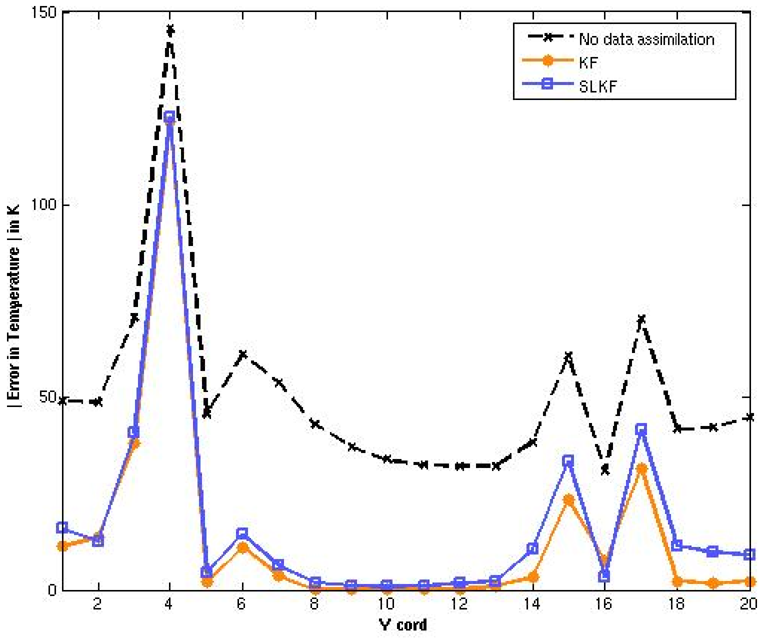
Data is injected into cells indicated by □

2D-Heat Conduction Example

- Compare temperature profile determined by the SLKF with the truth model across cross sections



Temperature distribution of the plate at t=50 s



Absolute value of the error in temperature profile along X=6 at t=50 s

Correlated Process and Measurement Noise

Cross Correlation

- When w_k and v_k are correlated, let

$$\mathcal{E}\left[\begin{bmatrix} w_k \\ v_k \end{bmatrix} \begin{bmatrix} w_k^\top & v_k^\top \end{bmatrix}\right] = \begin{bmatrix} Q_k & S_k \\ S_k^\top & R_k \end{bmatrix}$$

  – S_k is the cross correlation matrix

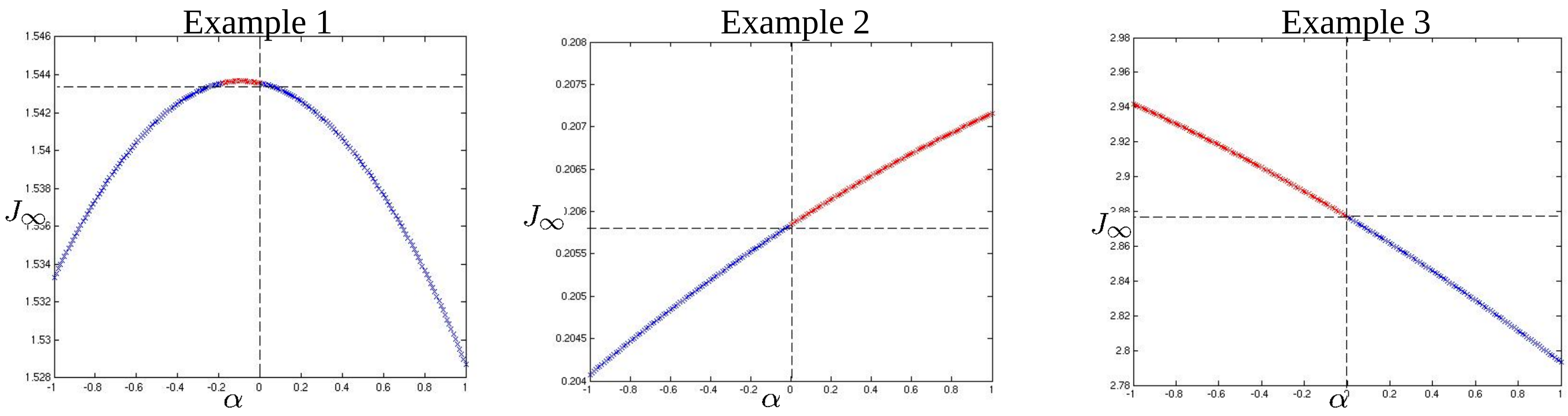
- Filter equations

$$\begin{aligned}
\hat{x}_{k+1} &= A_k \hat{x}_k + B_k u_k + K_k (y_k - \hat{y}_k) \\
\hat{y}_k &= C_k \hat{x}_k \\
K_k &= (A_k P_k C_k^\top + S_k)(C_k P_k C_k^\top + R_k)^{-1} \\
P_{k+1} &= A_k P_k A_k + Q_k - \underbrace{(A_k P_k C_k^\top + S_k)(C_k P_k C_k^\top + R_k)^{-1}(A_k P_k C_k^\top + S_k)^\top}_{\text{Uncertainty reduction due to filtering}}
\end{aligned}$$

  – Reduces uncertainty even if $C_k = 0$, i.e. $y_k = v_k$

Cross Correlation Example

- LTI example : $n = 2,\ w_k \in \mathbb{R}^2,\ v_k \in \mathbb{R}^1,\ S_0 \in \mathbb{R}^{2\times 1}$
- Set $S = \alpha S_0$
  - α varies from -1 to 1
- Compare final cost $J_\infty \triangleq \lim_{k\to\infty} \mathrm{tr} P_k$



- Correlation helps

- Correlation hurts

Unknown Noise Covariances

Unknown Noise Covariances

- True noise covariances Q and R are unknown
- Assume we use $\widehat{Q}$ and $\widehat{R}$
- Error dynamics

$$e_{k+1} = (A - \widehat{K}_k C)e_k + w_k - \widehat{K}_k v_k$$

$$\widehat{K}_k = A\widehat{P}_k C\left(C\widehat{P}_k C^\top + \widehat{R}\right)^{-1}$$

$$\widehat{P}_{k+1} = A\widehat{P}_k A^\top - A\widehat{P}_k C^\top (C\widehat{P}_k C^\top + \widehat{R})^{-1} C\widehat{P}_k A^\top + \widehat{Q}$$

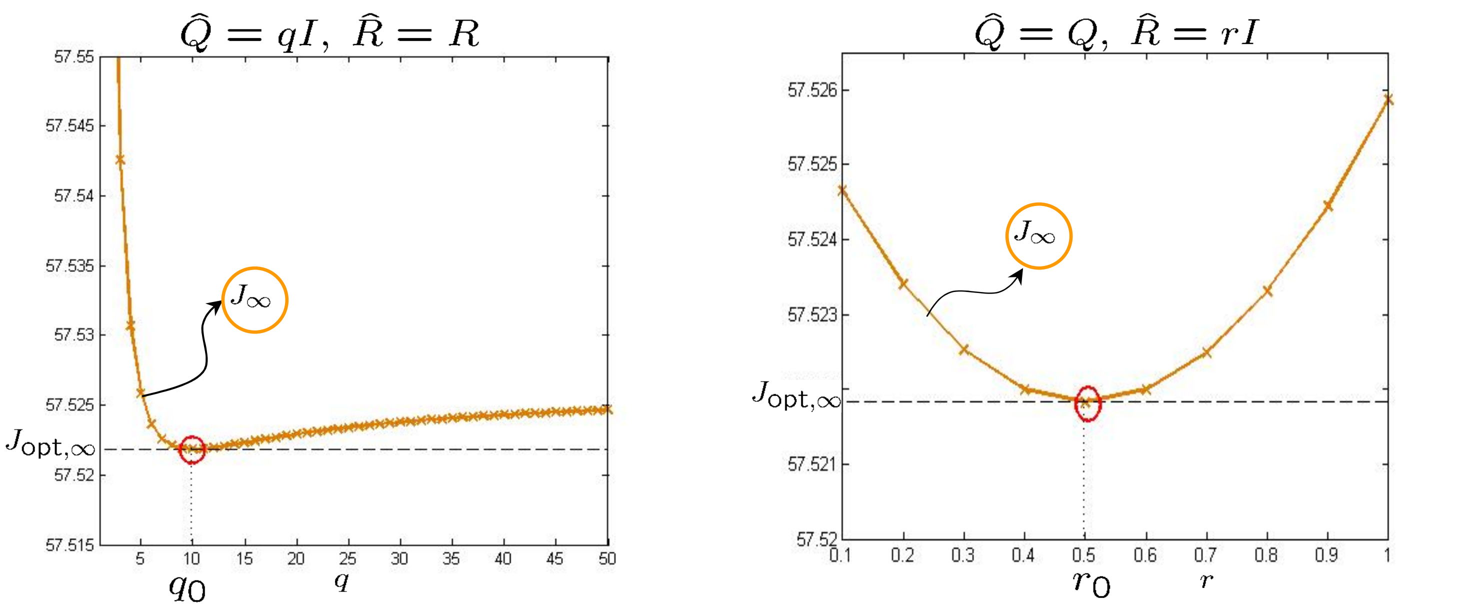
- The estimates are not the optimal estimates
- $\widehat{P}_k$ is not the error covariance (*pseudo error covariance*)
- Actual error covariance satisfies

$$P_{k+1} = (A - \widehat{K}_k C)P_k (A - \widehat{K}_k C)^\top + Q + \widehat{K}_k R \widehat{K}_k^\top$$

- If $(A, \widehat{Q})$ is stabilizable, the filter converges to an asymptotically stable observer

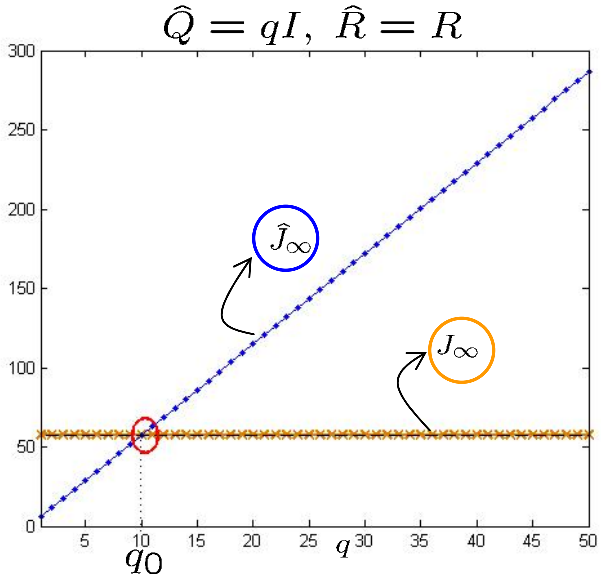
Incorrect Noise Covariances

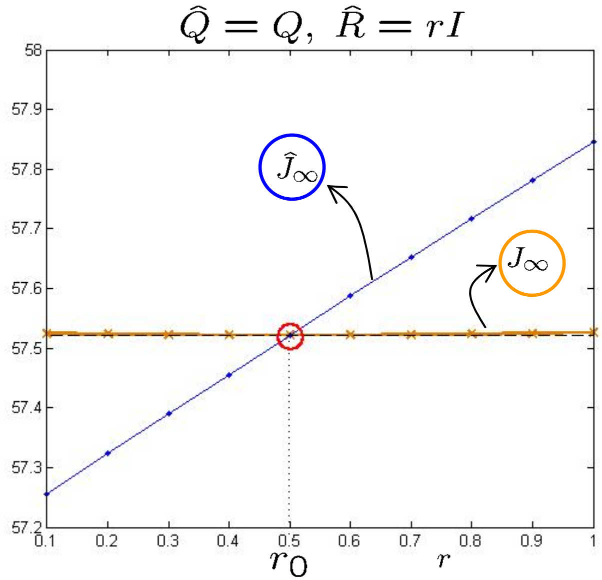
- If $\hat{Q} \neq Q, \ \hat{R} \neq R$ then $\underbrace{P_{\text{opt},k}}_{\text{Optimal error covariance}} \leqslant \underbrace{P_k}_{\text{Actual error covariance}}$
- LTI system example
  - $Q = q_0 I, \ R = r_0 I$
- Actual cost $J_\infty \triangleq \lim_{k\to\infty} \text{tr} P_k$



Incorrect Noise Covariances

- Pseudo cost $\hat{J}_\infty \triangleq \lim_{k\to\infty} \mathrm{tr}\hat{P}_k$





- If $Q \leqslant \hat{Q}$ and $R \leqslant \hat{R}$ then $\underbrace{P_k}_{\text{Actual error covariance}} \leqslant \underbrace{\hat{P}_k}_{\text{Pseudo error covariance}}$ *(Heffes and Nishimura)*
  - Provides an upper bound for the worst case performance

Unknown Initial Condition Statistics

Incorrect Initial Estimate

- Assume mean $\bar{x}_0$ of the random variable x_0 is unknown
- Assume we use $\hat{x}_0 \neq \bar{x}_0$
- Mean-error dynamics : $\bar{e}_k \triangleq \mathcal{E}[e_k]$

$$\bar{e}_k = \left[\prod_{i=0}^{k} (A - \hat{K}_k C)\right] \bar{e}_0$$

$$\hat{K}_k = A\hat{P}_k C \left(C\hat{P}_k C^\top + R\right)^{-1}$$

$$\hat{P}_{k+1} = A\hat{P}_k A^\top - A\hat{P}_k C^\top (C\hat{P}_k C^\top + R)^{-1} C\hat{P}_k A^\top + Q,\ \hat{P}_0 = \text{var}(x_0)$$

Pseudo error covariance

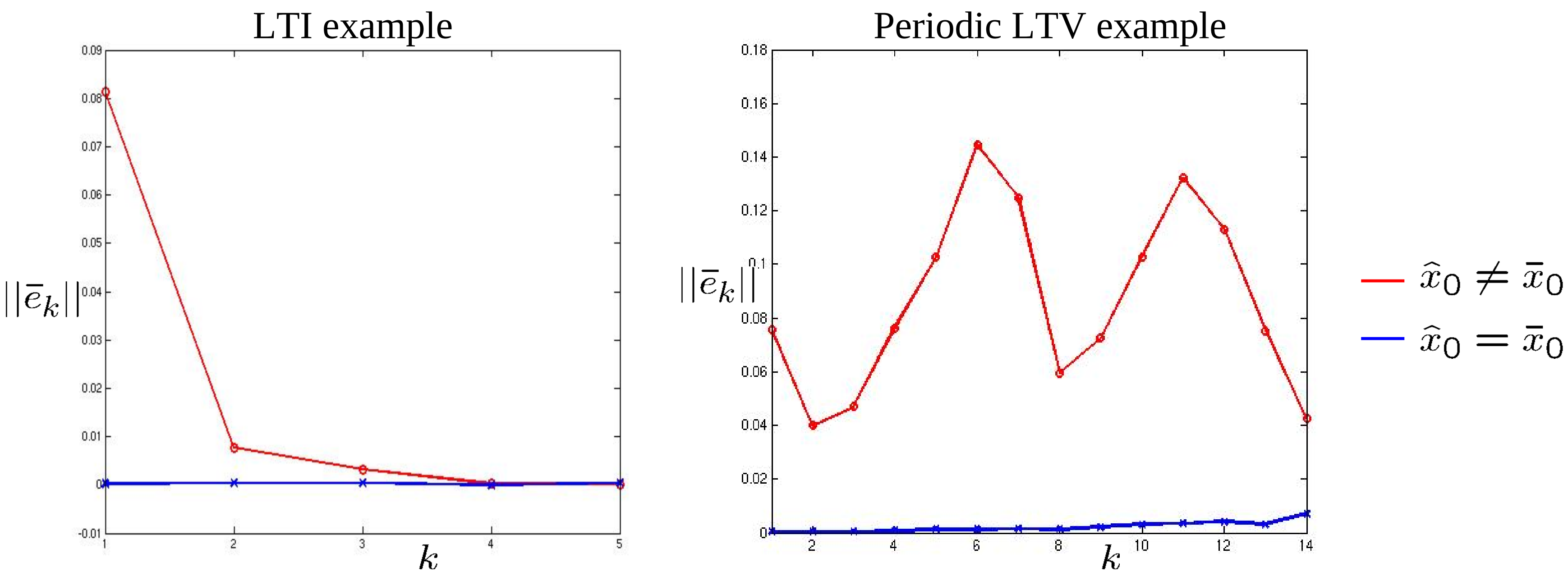
known

- True error covariance satisfies $P_k = \mathcal{E}[e_k e_k^\top] - \bar{e}_k \bar{e}_k^\top$
- Estimate $\hat{x}_k$ of x_k will be biased, i.e. $\bar{e}_k \neq 0$
- If the filter is stable $\lim_{k\to\infty} \bar{e}_k = 0$

Incorrect Initial Estimate

- Monte-Carlo Simulation (Sample size N = 10000)
- Compare $\bar{e}_k$
  - $\hat{x}_0 = \bar{x}_0$ versus $\hat{x}_0 \neq \bar{x}_0$

$$\bar{e}_k = \frac{1}{N} \sum_{i=1}^{N} (e_k)_i$$



- In the LTI case - $A - K_k C$ is asymptotically stable : $\bar{e}_k \to 0$ even when $\hat{x}_0 \neq \bar{x}_0$

Incorrect Initial State Covariance

- Assume $\mathrm{var}(x_0)$ is the true variance of x_0 and unknown
- Assume we use the incorrect initial covariance
- Error dynamics

$$e_{k+1} = (A - \widehat{K}_k C)e_k + w_k - \widehat{K}_k v_k$$

$$\widehat{K}_k = A\widehat{P}_k C \left(C\widehat{P}_k C^\top + R\right)^{-1}$$

$$\widehat{P}_{k+1} = A\widehat{P}_k A^\top - A\widehat{P}_k C^\top (C\widehat{P}_k C^\top + R)^{-1} C\widehat{P}_k A^\top + Q, \quad \widehat{P}_0 \neq \mathrm{var}(x_0)$$

Pseudo error covariance

- Estimate $\widehat{x}_k$ of x_k is unbiased

Not optimal

- If the filter is stable

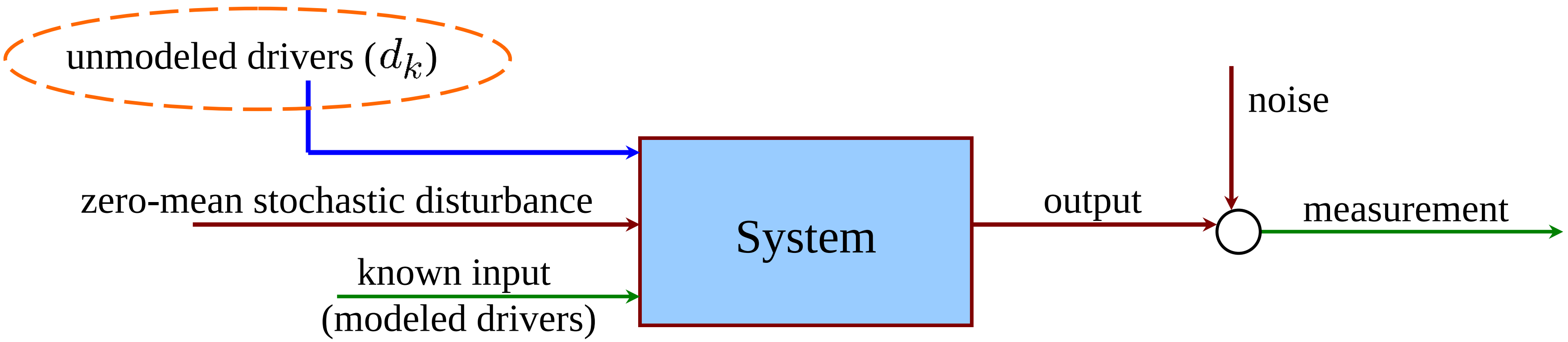
Actual error covariance

$$\lim_{k\to\infty} \widehat{P}_k = \lim_{k\to\infty} P_k = P$$

  – the estimates will converge to the optimal estimates

Unmodeled Drivers

Unmodeled Drivers



- Unmodeled driver can be deterministic or stochastic

Standard Kalman Filter

- Estimator dynamics

$$e_{k+1} = (A_k - \widehat{K}_k C)e_k + H_k \underbrace{d_k}_{\text{unknown}} + w_k - \widehat{K}_k v_k$$

$$\widehat{K}_k = A_k \widehat{P}_k C \left(C_k \widehat{P}_k C_k^\top + R_k \right)^{-1}$$

$$\widehat{P}_{k+1} = A_k \widehat{P}_k A_k^\top - A_k \widehat{P}_k C_k^\top (C_k \widehat{P}_k C_k^\top + R_k)^{-1} C_k \widehat{P}_k A_k^\top + Q_k$$

  - Estimates are not optimal
  - Estimates are biased due to d_k
  - $\widehat{P}_k$ is not the actual error covariance

Pseudo error covariance

Problem Formulation

- System

$$
\begin{aligned}
x_{k+1} &= A_k x_k + B_k u_k + H_k d_k + w_k \\
y_k &= C_k x_k + v_k
\end{aligned}
$$

- $A_k,\ B_k,\ C_k,\ \boxed{H_k}$ are known

- Signals $u_k,\ y_k$ are measured

- Signal $d_k \in \mathbb{R}^p$ is unknown and arbitrary

- Obtain unbiased estimates of states $x_k \in \mathbb{R}^n$
- Estimate the unknown signal $d_k \in \mathbb{R}^p$

Unbiasedness

- Two-step filter

$$\begin{aligned} \hat{x}_k^{\mathsf{da}} &= \hat{x}_k^{\mathsf{f}} + K_k(y_k - C_k \hat{x}_k^{\mathsf{f}}) \\ \hat{x}_{k+1}^{\mathsf{f}} &= A_k \hat{x}_k^{\mathsf{da}} + B_k u_k \end{aligned}$$

- Unbiased if and only if (*Kitanidis 1987*)

$$\mathcal{E}[e_k^{\mathsf{da}}] = \mathcal{E}[x_k - \hat{x}_k^{\mathsf{da}}] = 0$$

$$\Leftrightarrow$$

$$(I - K_k C_k) H_{k-1} = 0$$

- Minimize $\mathrm{tr}(P_k^{\mathsf{da}}) \triangleq \mathrm{tr}\left(\mathcal{E}\left[e_k^{\mathsf{da}}(e_k^{\mathsf{da}})^{\mathsf{T}}\right]\right)$
  - Subject to constraint $(I - K_k C_k) H_{k-1} = 0$
  - Need $\mathrm{rank}(C_k H_{k-1}) = p$

Unbiased Minimum-variance Filter

- Define

$$V_k \triangleq C_k H_{k-1} \qquad \tilde{R}_k \triangleq C_k P_k^{\mathsf{f}} C_k^{\mathsf{T}} + R_k$$
$$F_k \triangleq P_k^{\mathsf{f}} C_k^{\mathsf{T}} \qquad \Pi_k \triangleq (V_{k+1}^{\mathsf{T}} \tilde{R}_{k+1}^{-1} V_{k+1})^{-1} V_{k+1}^{\mathsf{T}} \tilde{R}_{k+1}^{-1}$$

- Optimal filter gain

$$K_k = H_{k-1}\Pi_{k-1} + F_k \tilde{R}_k^{-1}(I - V_k \Pi_{k-1})$$

- Covariance update

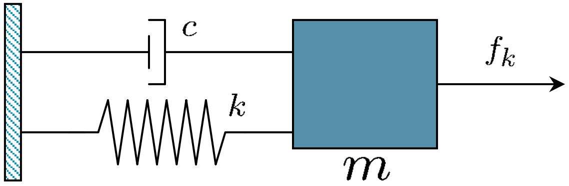
$$P_k^{\mathsf{da}} = K_k R_k K_k^{\mathsf{T}} + (I - K_k C_k) P_k^{\mathsf{f}} (I - K_k C_k)^{\mathsf{T}}$$
$$P_{k+1}^{\mathsf{f}} = A_k P_k^{\mathsf{da}} A_k^{\mathsf{T}} + Q_k$$

- Reduces to Kalman filter when $H_k = 0$
- Unbiased estimate $\hat{d}_k$ of d_k obtained as

$$\hat{d}_k = H_k^{\dagger} L_{k+1}(y_{k+1} - C_{k+1}\hat{x}_{k+1}^{\mathsf{f}})$$

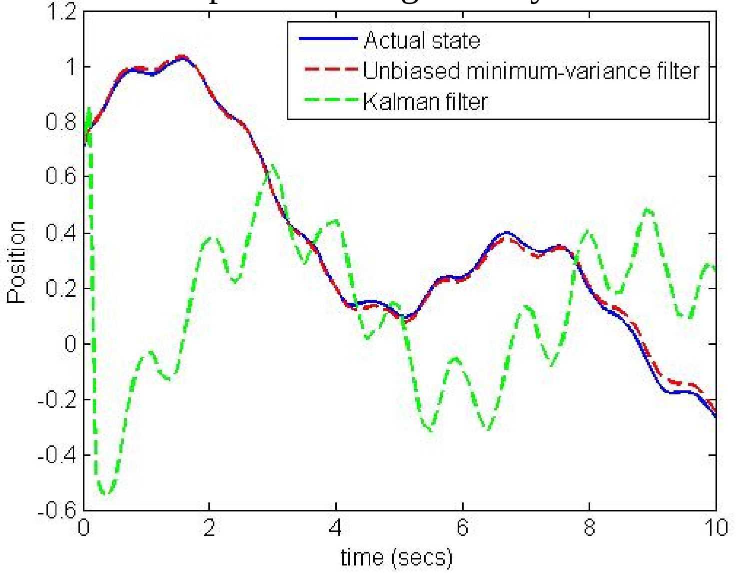
Estimation with unknown inputs

- Mass spring damper

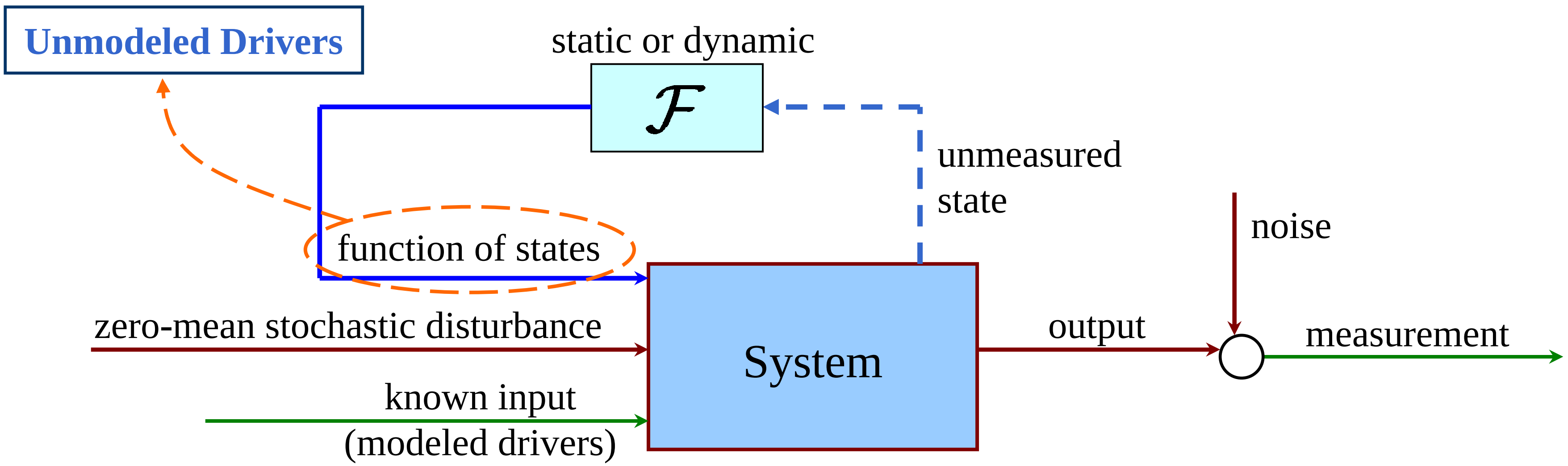


$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} x_{1,k} + t_s x_{2,k} \\ x_{2,k} + t_s(-\frac{k}{m}x_{1,k}^2 - \frac{c}{m}x_{2,k}) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} f_k$$

Unknown force



Unknown Feedback Signal



- Unmodeled drivers can be feedback signals
- Estimates of states and unknown signal are still unbiased !!

Example

- Discretized Van der Pol Oscillator

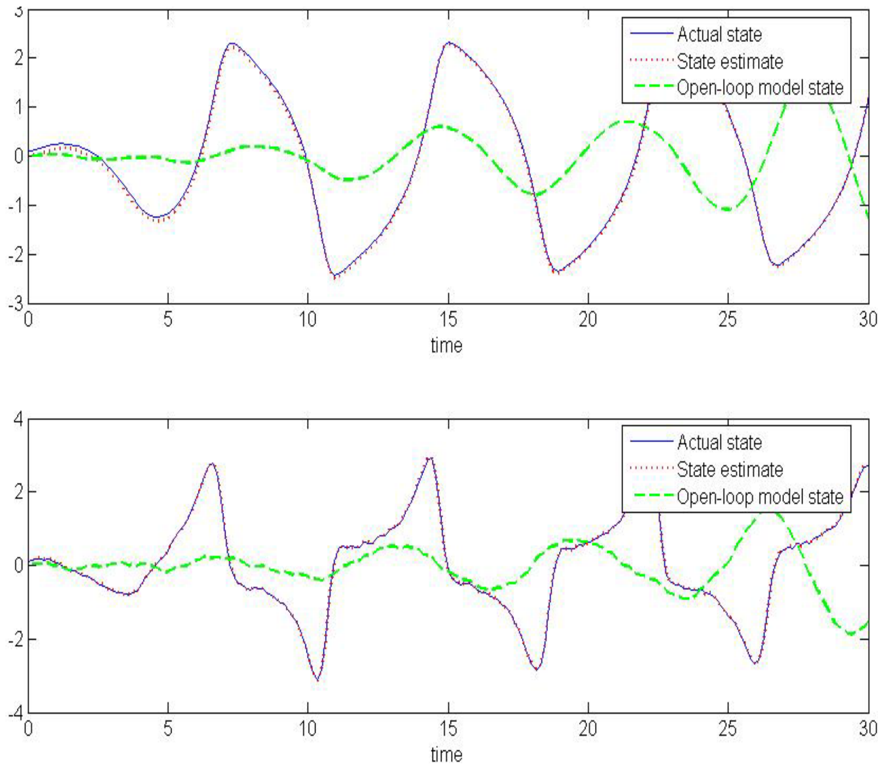
$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} x_{1,k} + t_S x_{2,k} \\ x_{2,k} + t_S[\underbrace{(1 - x_{1,k}^2)x_{2,k}}_{\text{unknown nonlinearity} \sim d_k} - x_{1,k} + u_k] \end{bmatrix}$$

$$y_k = x_{2,k}$$

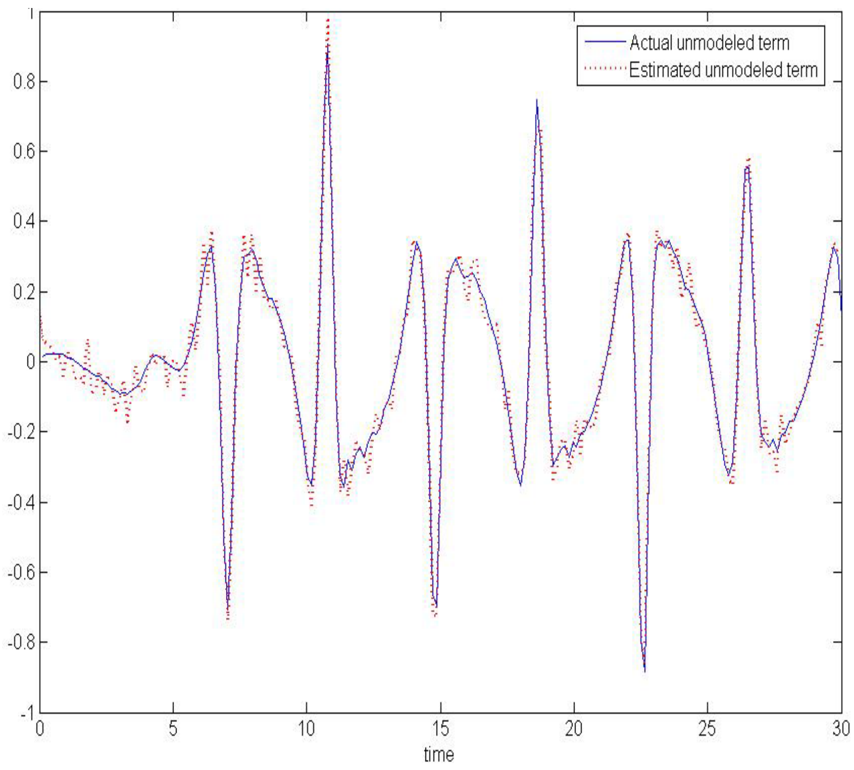
  - t_S is the sample interval
  - $H_k = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and the linear part of the dynamics is known
  - $d_k = t_S(1 - x_{1,k}^2)x_{2,k}$ is the unknown (unmodeled) signal

Example

State estimates



Estimate $\hat{d}_k$ of unmodeled signal d_k



Nonlinear Systems

Estimators for Nonlinear Systems

- System dynamics

$$
\begin{aligned}
x_{k+1} &= f(x_k, u_k, k) + w_k \\
y_k &= h(x_k, k) + v_k
\end{aligned}
$$

- Estimator dynamics
  - One-step

$$
\begin{aligned}
\hat{x}_{k+1} &= \underbrace{f(\hat{x}_k, u_k, k)}_{\text{nonlinear dynamics}} + \underbrace{K_k(y_k - \hat{y}_k)}_{\text{innovation}} \\
\hat{y}_k &= h(\hat{x}_k, k)
\end{aligned}
$$

  - Two-step

$$
\begin{aligned}
x^{\mathsf{f}}_{k+1} &= \underbrace{f(x^{\mathsf{da}}_k, u_k, k)}_{\text{nonlinear dynamics}} \\
x^{\mathsf{da}}_k &= x^{\mathsf{f}}_k + \underbrace{K_k(y_k - y^{\mathsf{f}}_k)}_{\text{innovation}} \\
y^{\mathsf{f}}_k &= h(x^{\mathsf{f}}_k, k)
\end{aligned}
$$

Nonlinear Filter Theory

- One-step and two-step estimators not equivalent
- x_k may not be Gaussian even if $w_k,\ v_k$ and x_0 are Gaussian
- For continuous-time systems, the probability density function of x is governed by the Fokker-Planck partial differential equation

$$\frac{\partial}{\partial t}p(x,t|x_0,t_0) = \frac{1}{2}\frac{\partial^2}{\partial x^2}[\sigma^2(x,t)p(x,t|x_0,t_0)] - \frac{\partial}{\partial x}[m(x,t)p(x,t|x_0,t_0)]$$

  - Scalar case
  - $p(\cdot)$ = probability density function of x
  - $\sigma(\cdot)$ and $m(\cdot)$ depend on the nonlinear function $f(\cdot)$
  - Difficult to propagate actual covariance P_k

Nonlinear Filter Theory

- Optimal filters for nonlinear systems are usually infinite dimensional
  - Finite dimensional optimal filters exist for a limited class of nonlinear systems *(Daum)*

- Ad hoc idea : Use classical linear Kalman filter gain expression

$$K_k = A_k P_k C_k^\top \left(C_k P_k C_k^\top + R_k\right)^{-1}$$

- P_k is a pseudo error covariance

Estimation with Pseudo Covariance

- Extended Kalman Filter (XKF)

- State-Dependent Riccati Equation filter (SDRE)

- Particle filters (Monte Carlo Technique)
  - Unscented Kalman filter
  - Simplex Kalman filter
  - Ensemble Kalman Filter

Extended Kalman Filter

- Set $A_k = \underbrace{\frac{\partial f(x,u,k)}{\partial x}}_{\text{Jacobian}}\Big|_{x=\hat{x}_k}, \quad C_k = \frac{\partial h(x,k)}{\partial x}\Big|_{x=\hat{x}_k}$

- One-step estimator dynamics

$$P_{k+1} = A_k P_k A_k^\top - A_k P_k C_k^\top (C_k P_k C_k^\top + R_k)^{-1} C_k P_k A_k^\top + Q_k$$

$$K_k = A_k P_k C_k^\top \left(C_k P_k C_k^\top + R_k\right)^{-1}$$

$$\hat{x}_{k+1} = f(\hat{x}_k, u_k, k) + K_k (y_k - \hat{y}_k)$$

$$\hat{y}_k = h(\hat{x}_k, k)$$

- P_k is the pseudo error covariance
  - P_k is not the actual covariance of the error

XKF Properties

- Filter gain K_k depends on the state estimate $\widehat{x}_k$
  - Filter gains cannot be evaluated offline !
- Estimate $\widehat{x}_k$ of x_k may be biased even if $\widehat{x}_0 = \bar{x}_0$
- Stability of the filter cannot be guaranteed
- We consider the use of XKF for
  - Satellite orbit estimation
  - Data assimilation in one-dimensional hydrodynamic flow

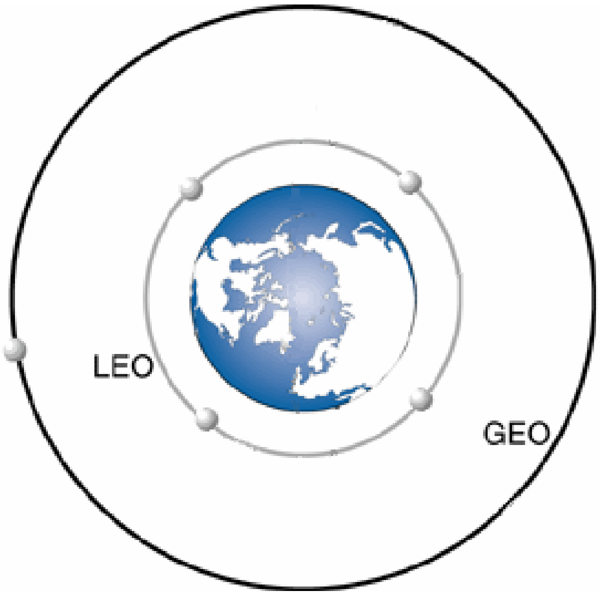
Satellite Orbit Estimation

- Problem: Track geosynchronous satellite with 4 observing satellites in low-Earth orbit

- Use Sampled-Data Extended Kalman Filter
  - Few sensors (range-only)
  - Time-sparse measurements

- Evaluate tradeoffs
  - Acquisition time, estimation accuracy

    versus
  - Measurement sample rate



Satellite Equations of Motion

- Orbiting Spacecraft Equations of Motion

$$\ddot{\vec{r}} = \frac{\mu}{r^3}\vec{r} + \vec{w}, \; r \triangleq \|\vec{r}\| = \sqrt{X^2 + Y^2 + Z^2}$$

- Measurement Model
  - Range data from l satellites at time $t = kt_s$

$$y_k = h(X(kt_s), Y(kt_s), Z(kt_s)) = \begin{bmatrix} h_1(X, Y, Z, X_1, Y_1, Z_1) \\ \vdots \\ h_l(X, Y, Z, X_l, Y_l, Z_l) \end{bmatrix} + v_k$$

$$h_i(X, Y, Z, X_i, Y_i, Z_i) = \sqrt{(X - X_i)^2 + (Y - Y_i)^2 + (Z - Z_i)^2}$$

  - Earth Blockage
    - Measurement is unavailable when line-of-sight between i^{th} observing satellite and target is blocked by the Earth

Sampled-Data XKF

- Measurements available every t_s seconds
  - Forecast Step (No data available): $t \in [kt_s, (k+1)t_s]$

$$\dot{\hat{x}}(t) = f(\hat{x}(t)), \;\; \hat{x}(kt_s) = \hat{x}(k+)$$

$$\dot{P}(t) = A(t)P(t) + P(t)A(t) + Q, \;\; P(kt_s) = P(k+)$$

Pseudo error covariance

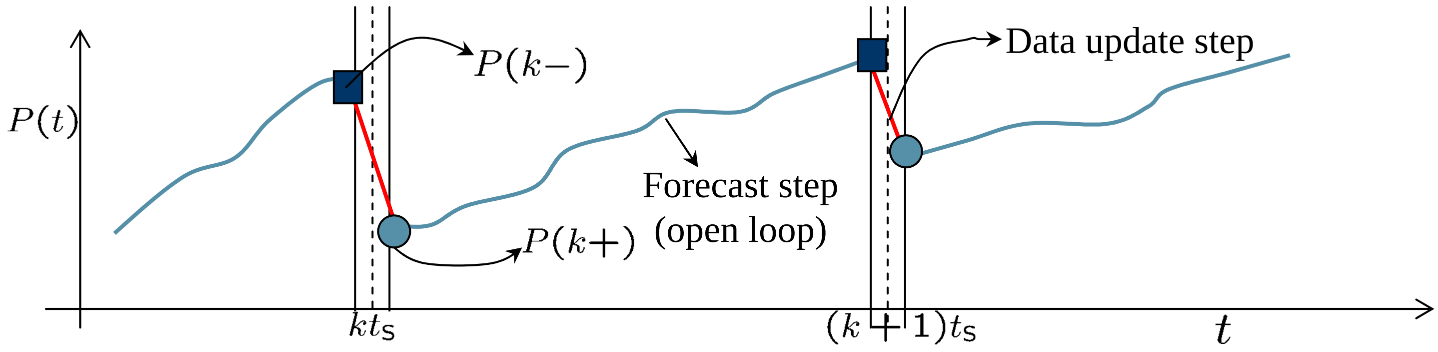
$$A(t) \triangleq f'(\hat{x}(t))$$

  - Data-Update Step

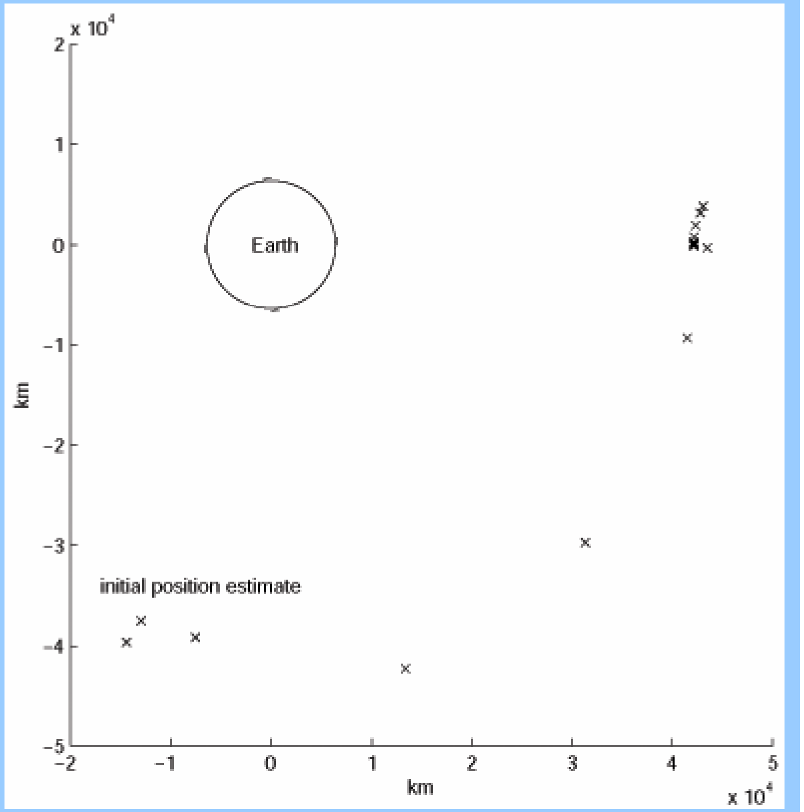
$$K_k = P(k-)C_k^{\top}(C_k P(k-)C_k^{\top} + R)^{-1}, \; C_k \triangleq h'(\hat{x}(k-))$$

$$\hat{x}(k+) = \hat{x}(k-) + K_k(y_k - \hat{d}_k)$$

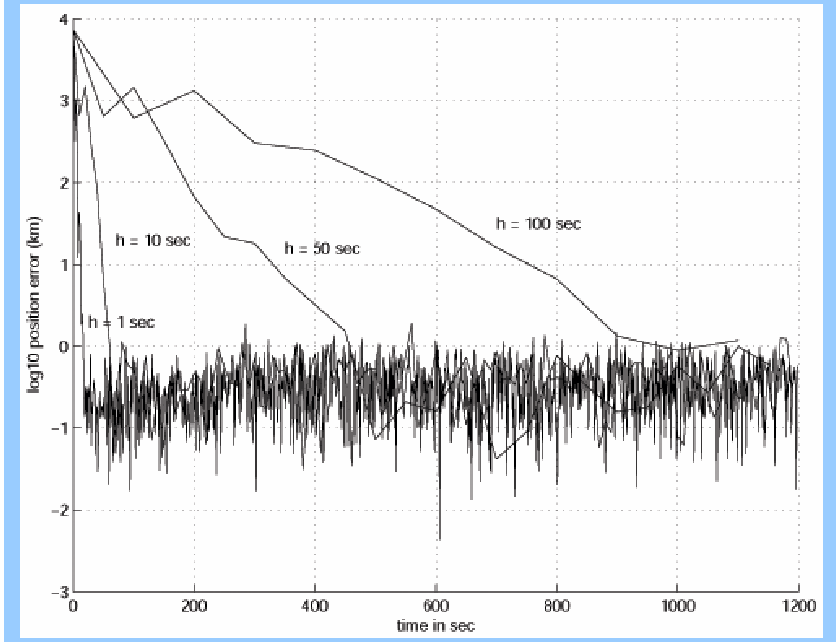
$$P(k+) = (I - K_k C_k)P(k-)$$



Target Acquisition

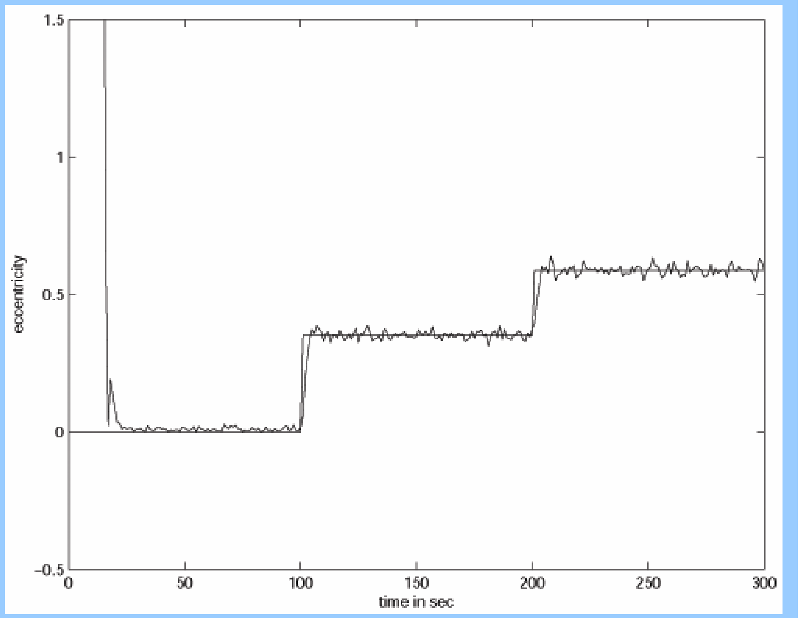


•Initial True Anomaly Error: 110°
•Sample Interval: 1s
•Meas. Standard Deviation: 0.1km

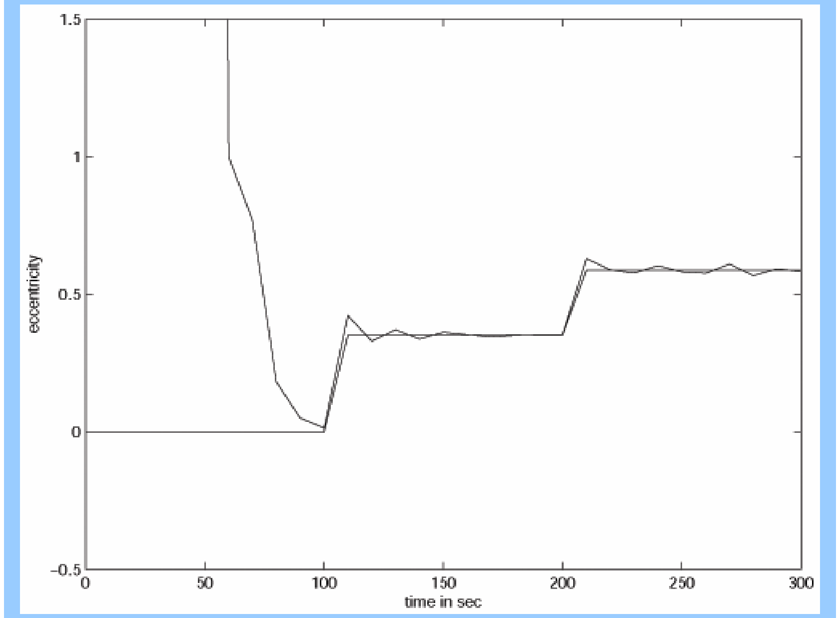


•Sample Intervals: 1s, 10s, 50s, 100s
•Meas. Standard Deviation: 0.1km

Eccentricity Estimation

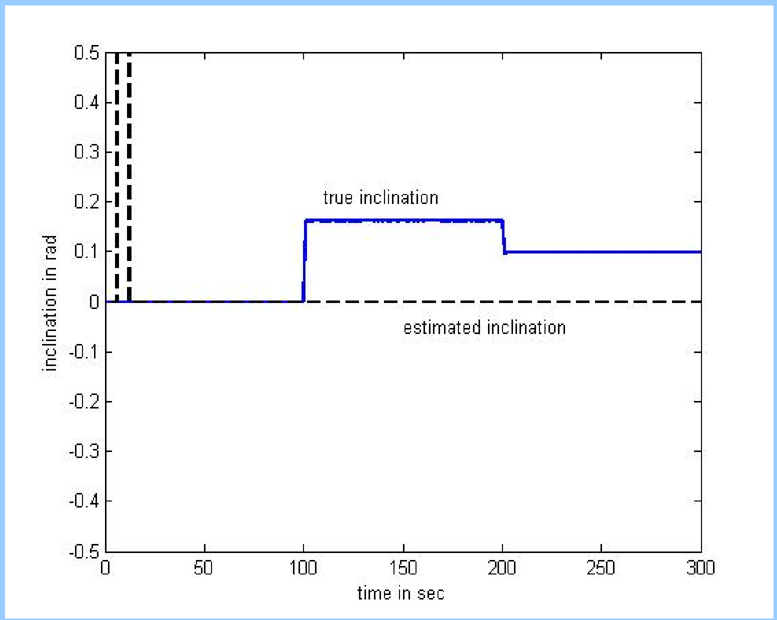


•Sample Interval = 1s
•Meas. Standard Deviation: 0.01km
•Target performs a 1s burn at t=100s and t=200s



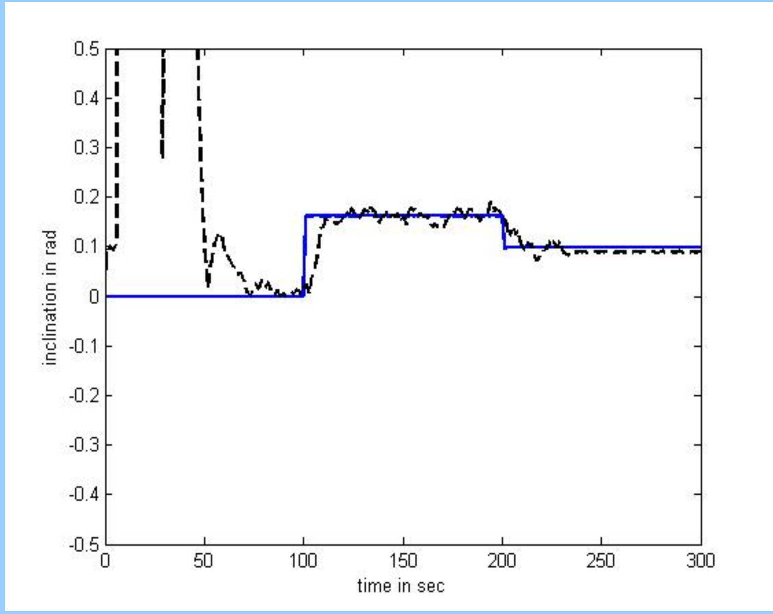
•Sample Interval = 10s
•Meas. Standard Deviation: 0.01km
•Target performs a 1s burn at t=100s and t=200s

Inclination Estimation



- All observing satellites in equatorial orbit
- Target performs a 1s burn at t=100s and t=200s
- Lack of observability

- Change inclination of two observing satellites (i = 0.1rad, i = -0.2rad)
- Sample Interval = 1s
- Successfully track inclination change



Ideal Hydrodynamic Equations

- Consider inviscid adiabatic flow along a 1-D channel
  - Flow is governed by Euler's equation and continuity equation in conservative form

$$\frac{\partial U}{\partial t} + \nabla \cdot \mathcal{F} = 0$$

- U is the vector of conserved quantities

$$U = \left[\begin{array}{ccc} \varrho & \varrho v & \mathcal{E} \end{array} \right]^{\mathsf{T}}$$

- $\mathcal{F}$ is the flux dyad

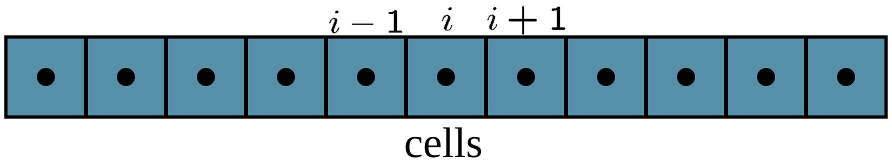
$$\mathcal{F} = \left[\begin{array}{c} \varrho v \\ \varrho v^2 + p \\ v(\mathcal{E} + p) \end{array} \right]$$

$$\begin{aligned} \varrho &= \text{density} \\ v &= \text{velocity} \\ \mathcal{E} &= \text{energy} \\ p &= \text{pressure} \\ \gamma &= \text{specific heat ratio} \end{aligned}$$

$$\mathcal{E} = \frac{1}{1}\varrho v^2 + \frac{p}{\gamma - 1}$$

Finite Volume Model

- Discretize space into cells
  - Grid size depends on the required resolution
  - Number of cells can vary with time
  - 1-D grid
    - $U^{[i]}(k) =$ value of U at the center of the i^{th} cell at time step k



- Use second-order Rusanov scheme to determine flow variables in each discretized cell

$$U^{[i]}(k+1) = U^{[i]}(k) - \frac{t_{\text{s}}}{\delta x}[F_{\text{Rus}}^{[i]}(k) - F_{\text{Rus}}^{[i-1]}(k)]$$

- $F_{\text{Rus}}^{[i]}$ is the second-order Rusanov flux
  - determined using $U^{[i+n]}$, $n = -2, -1, \ldots, 2$
  - depend on the slope limiter (*minmod, MC*)
  - $\frac{t_{\text{s}}}{\delta x}$ satisfy the CFL stability condition

Discrete-Time Dynamic Model

- State contains values of all conserved quantities in all cells

$$x_k = \left[\, (U^{[1]}(k))^{\mathsf{T}} \;\; (U^{[2]}(k))^{\mathsf{T}} \;\; \ldots \;\; (U^{[n]}(k))^{\mathsf{T}} \, \right]^{\mathsf{T}}$$

- High dimensional, highly nonlinear dynamics

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

  – $f(\cdot)$ depends on the order and scheme used in the finite volume MHD flow simulation

- Involves modeled and unmodeled drivers
  – u_k represents known boundary conditions
  – w_k represents uncertainty in boundary conditions and modeling errors

XKF for 1-D HD Flow

- Nondifferentiable nonlinearities are present in the finite volume dynamics
  - For example : *abs*, *sgn, min* and *max* functions
  - Jacobian not exist due nondifferentiable nonlinearities
  - Differentiable approximations can be constructed
    - For example : $|x| \approx \mathtt{atan}(x)x$
  - Alternatively, numerical approximations of the Jacobians can be used

State-Dependent Riccati Equation

- Express nonlinear dynamics as a frozen-time pseudo-linear difference equation

$$x_{k+1} = \mathcal{A}(x_k)x_k + \mathcal{B}(x_k, u_k, w_k)$$

- Set $A_k = \mathcal{A}(\widehat{x}_k)$ in the covariance update and filter gain expression
  - The parameterization $\mathcal{A}(x)$ is not unique
  - Example :

$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = f(x_k) = \begin{bmatrix} x_{1,k}x_{2,k} \\ x_{1,k}^2 x_{2,k} + x_{2,k} \end{bmatrix}$$

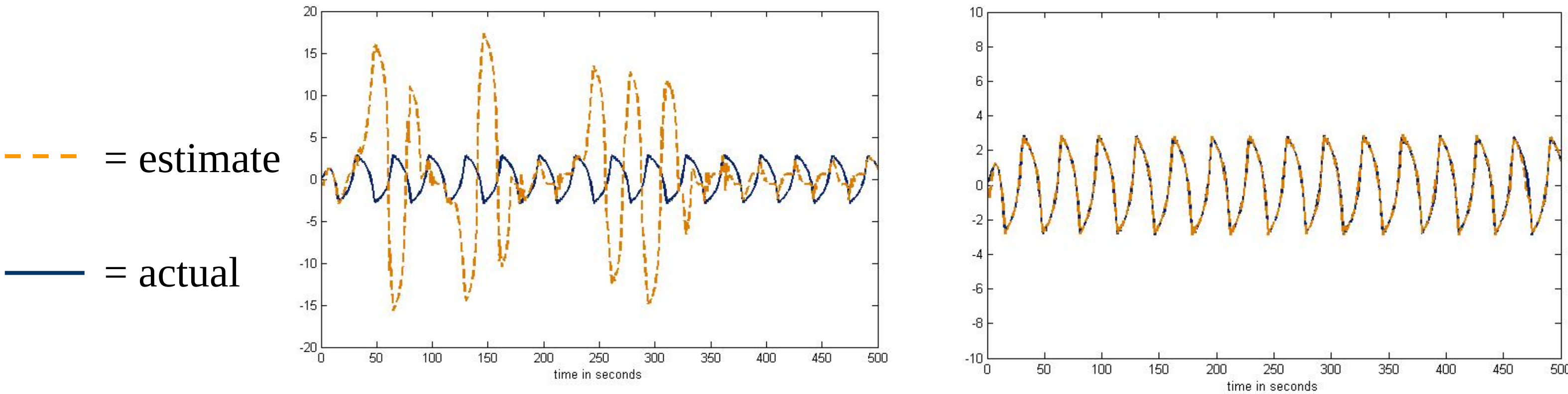
$$f(x_k) = \begin{bmatrix} x_{2,k} & 0 \\ 0 & x_{1,k}^2 + 1 \end{bmatrix} \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix}, \qquad f(x_k) = \begin{bmatrix} 0 & x_{1,k} \\ x_{1,k}x_{2,k} & 1 \end{bmatrix} \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix}$$

- Does not require the Jacobian !

State-Dependent Riccati Equation

- Performance depends on the parameterization $\mathcal{A}(x)$
  - Van der Pol Oscillator example
  - Use measurements of velocity to estimate position

Position estimates



$$A(x) = \begin{bmatrix} 1 & h \\ -t_s & 1 + t_s b(1 - x_1)^2 \end{bmatrix} \qquad A(x) = \begin{bmatrix} 1 & h \\ -t_s(1 + b x_1 x_2) & 1 + t_s b \end{bmatrix}$$

- Under certain conditions, can guarantee $\lim_{k \to \infty} [x_k - \widehat{x}_k] = 0$
  - Assumes a deterministic setting (Asymptotic observer)
  - These conditions are conservative

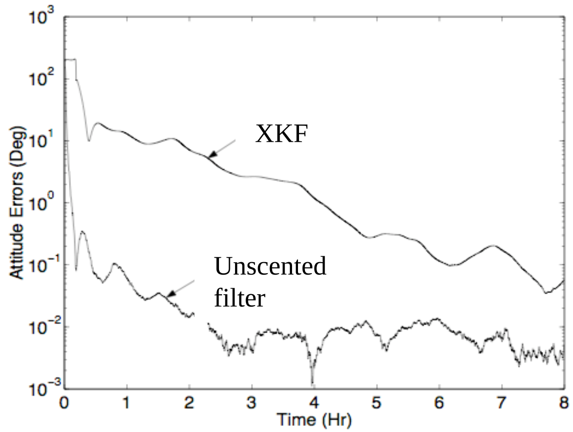
- **Open problem** : Finding optimal parameterizations

XKF versus SDRE-KF

XKF	SDRE-KF
The Jacobian of a particular nonlinear system is unique	Parameterization of a particular nonlinear system is **not unique**. Various parameterizations have to be determined and their performance evaluated.
Jacobian may have to be determined numerically due to the presence of nondifferentiable nonlinearities. Computationally intensive	Evaluating a parameterization is computationally less intensive compared to obtaining the Jacobian numerically
Knowledge of the dynamics is not necessary to obtain a numerical approximation of the Jacobian.	Knowledge of the exact dynamics is necessary to determine a parameterization

Particle Filter

- Run ensemble of estimators in parallel
- Compute ensemble estimates at every step
  - Motivation: The statistics of the ensemble members approximate that of the true state
- The “optimal” estimate is the average of the ensemble estimates
- Performs better than the XKF in certain applications
  - XKF retains only the first two terms in the Taylor series approximation of the error covariance
  - Particle filters retain higher order terms

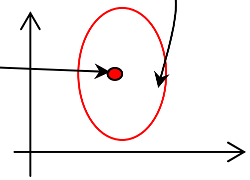


Spacecraft attitude estimation
(Crassidis and Markley, 2003)

Error in attitude estimates

Unscented Kalman Filter

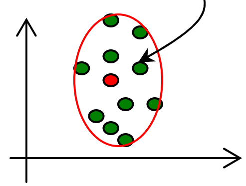
- Let $\hat{x}_k = \text{mean}(x_k)$, $\hat{P}_k = \text{var}(x_k)$



- Choose ensemble members $\hat{x}_k^i$ so that

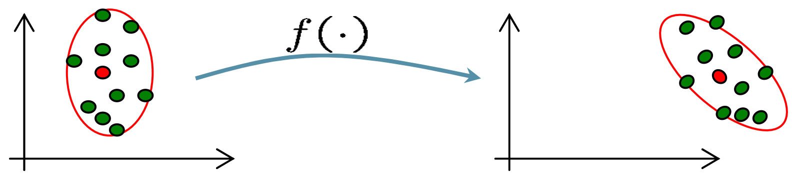
$$\text{var}_i(\hat{x}_k^i) = \hat{P}_k$$
$$\text{mean}_i(\hat{x}_k^i) = \hat{x}_k$$

(unscented transformation)



ensemble size = $2n + 1$

- Propagate the ensemble members through the nonlinearity

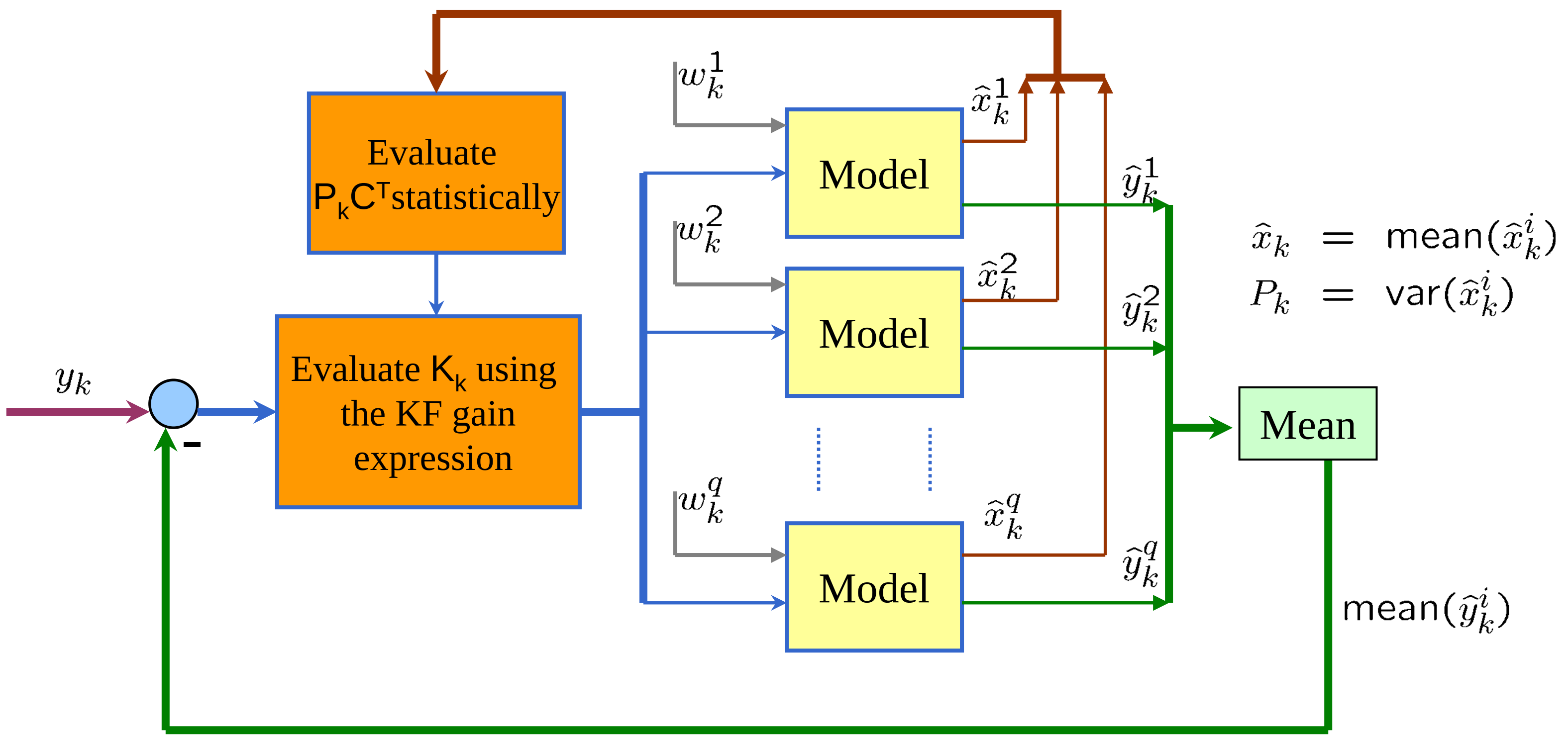


- Use the ensemble members to estimate the mean $\hat{x}_{k+1}$ and variance $\hat{P}_{k+1}$ of x_{k+1}

$$\hat{x}_{k+1} = \text{mean}\left[f(\hat{x}_k^i) + K_k(y_k - h(\hat{x}_k^i)\right]$$
$$\hat{P}_{k+1} = \text{var}\left[f(\hat{x}_k^i) + K_k(y_k - h(\hat{x}_k^i)\right]$$

Ensemble Kalman Filter

- Run ensemble of multiple estimators in parallel
  - Inject random disturbance into the ensembles (Monte Carlo)
- Initialize estimators with random initial conditions



Ensemble Kalman Filter

- Use estimates from the ensemble to approximate the error covariance at every time step

$$\Delta x_k^{\mathsf{f}} = \left[\, x_k^{\mathsf{f}_1} - \bar{x}_k^{\mathsf{f}} \quad \cdots \quad x_k^{\mathsf{f}_q} - \bar{x}_k^{\mathsf{f}} \,\right], \quad \bar{x}^{\mathsf{f}} = \frac{1}{q}\sum_{i=1}^{q} x_k^{\mathsf{f}_i}$$

$$\Delta y_k^{\mathsf{f}} = \left[\, y_k^{\mathsf{f}_1} - \bar{y}_k^{\mathsf{f}} \quad \cdots \quad y_k^{\mathsf{f}_q} - \bar{y}_k^{\mathsf{f}} \,\right], \quad \bar{y}_k^{\mathsf{f}} = \frac{1}{q}\sum_{i=1}^{l} y_k^{\mathsf{f}_i}$$

$$\boxed{P_k C_k^{\mathsf{T}} \sim \widehat{P}_{\mathsf{xy}_k}^{\mathsf{f}}} = \frac{1}{q-1}\Delta x_k^{\mathsf{f}}(\Delta y_k^{\mathsf{f}})^{\mathsf{T}}, \quad \boxed{C_k P_k C_k^{\mathsf{T}} + R_k \sim \widehat{P}_{\mathsf{yy}_k}^{\mathsf{f}}} = \frac{1}{q-1}\Delta y_k^{\mathsf{f}}(\Delta y_k^{\mathsf{f}})^{\mathsf{T}}$$

  - No error covariance update using the Riccati equation !
  - Number of operations $= \mathcal{O}(qn^2)$
  - n is the dimension of the system
  - q is the number of ensembles
- Data assimilation step

$$x_k^{\mathsf{da}_i} = x_k^{\mathsf{f}_i} + \widehat{K}_k\left(y_k - y_k^{\mathsf{f}_i}\right)$$

$$\boxed{\widehat{K}_k = \widehat{P}_{\mathsf{xy}_k}^{\mathsf{f}}(\widehat{P}_{\mathsf{yy}_k}^{\mathsf{f}})^{-1}}$$

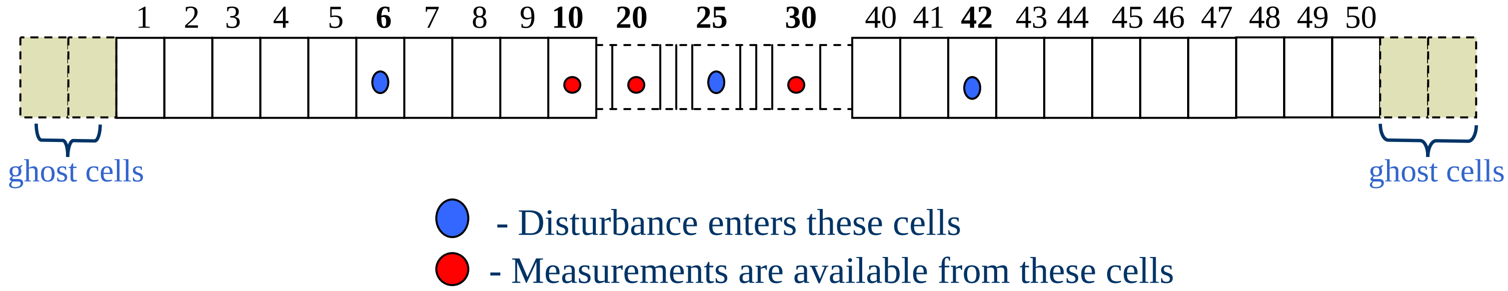
- The $n \times n$ error covariance $\widehat{P}_{\mathsf{xx}_k}^{\mathsf{f}}$ is never evaluated
  - Only the correlation $\widehat{P}_{\mathsf{xy}_k}^{\mathsf{f}}$ and $\widehat{P}_{\mathsf{yy}_k}^{\mathsf{f}}$ are evaluated

Ensemble Kalman Filter

- Computationally equivalent to running a collection of nonlinear simulations in parallel
  - Size of ensemble is critical !
    - Statistics of w_k and v_k must be accurately captured
- Extensive application to terrestrial weather prediction

Simulation : 1D- Hydrodynamics

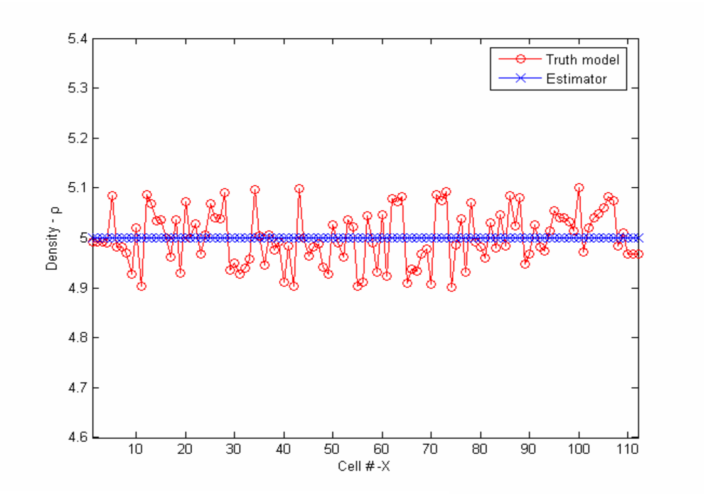
- Grid size = $N = 50$ cells

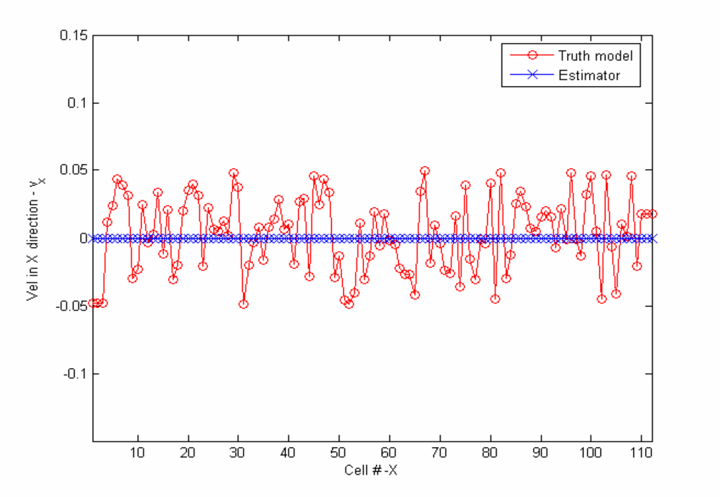


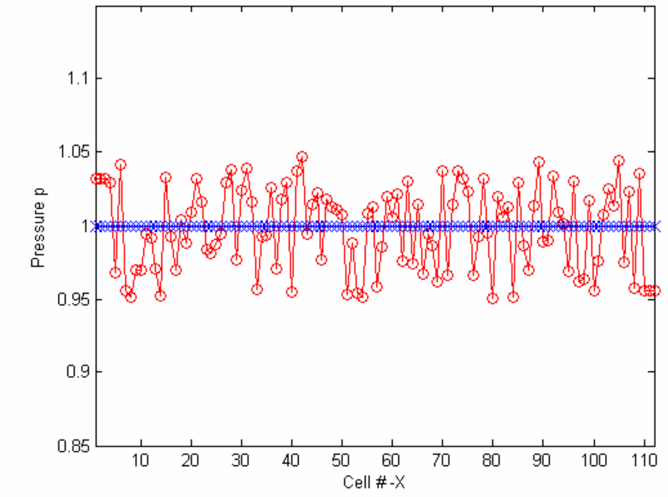
- Dimension of state vector $X = 3N$
- Measurements of density, velocity and pressure (corrupted by sensor noise) are available at some of the cells
- Boundary conditions are determined by the flow variables in the ghost cells
- Compare performance of different data assimilation techniques

Initial Condition

- Density, velocity and pressure distribution of the “truth” model and the estimator at $t = 0$

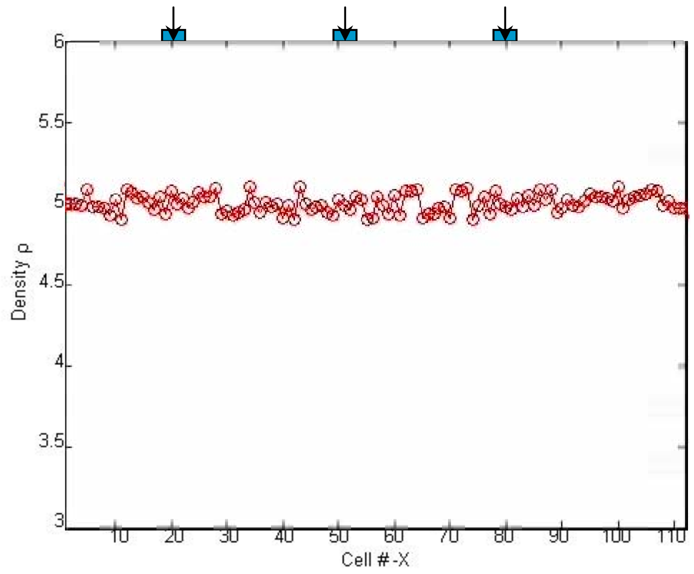






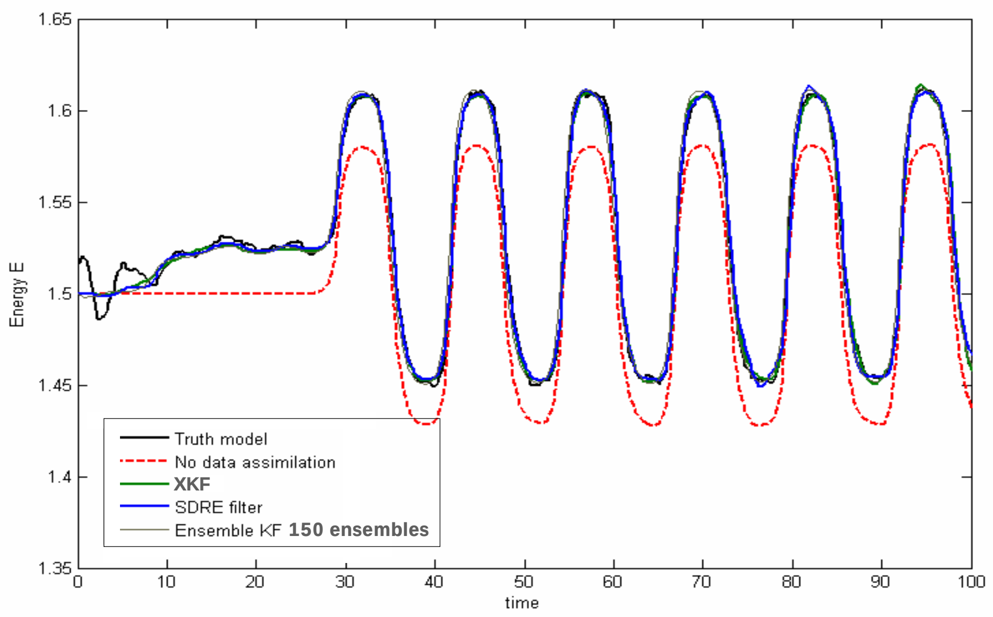
Flow Conditions

- Subsonic flow
- Boundary conditions
  - Left
    - Constant density and pressure
    - Sinusoidally varying velocity
  - Right
    - Floating boundary conditions
- Disturbance enter the cells indicated by



Simulation : 1D- HD

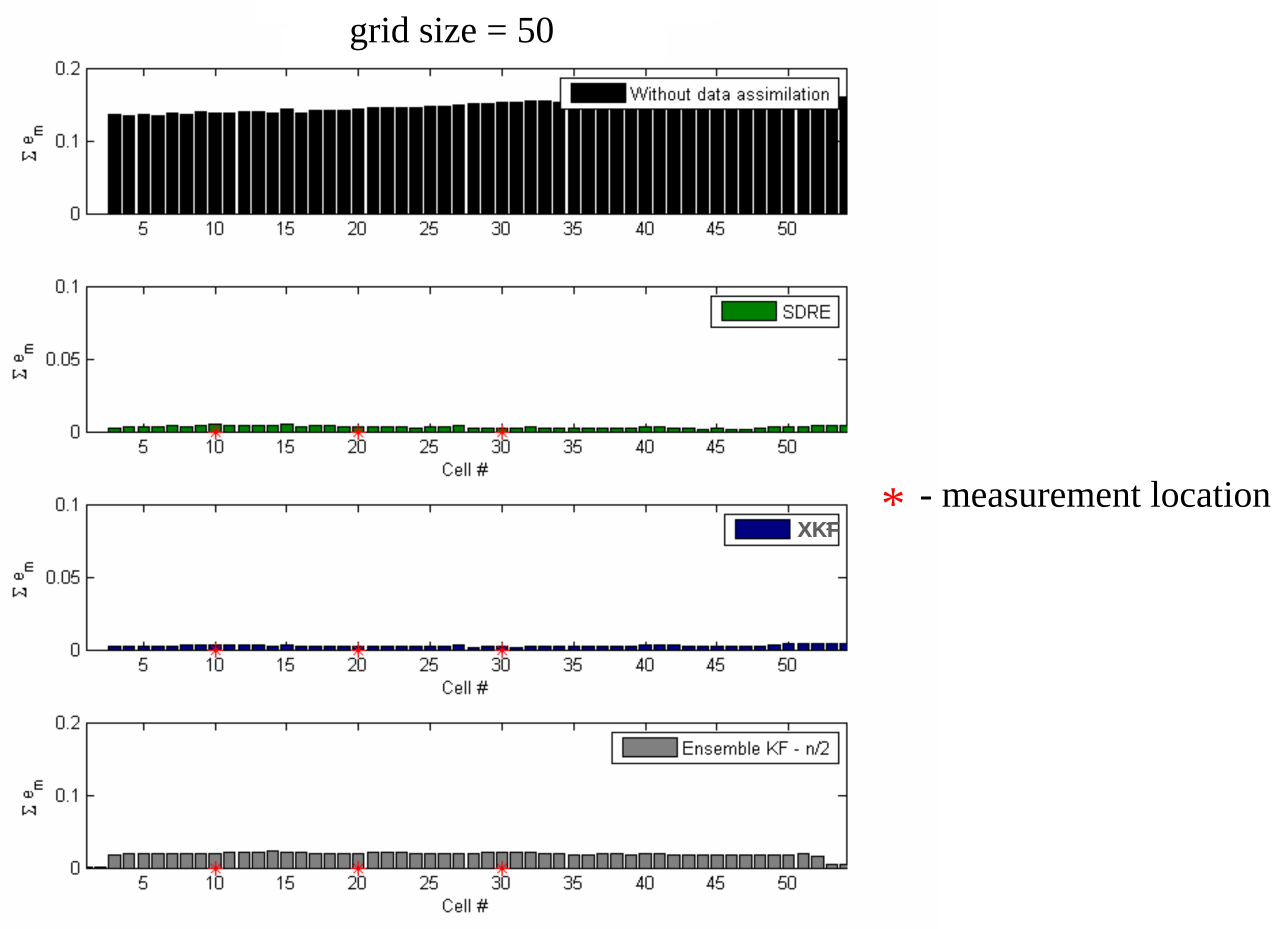
- Grid size = 50
- Measurement available at cells 10, 20, and 30
- Compare estimate of energy at cell 40



- Transient is due to the difference in the initial conditions between the “truth” model and the estimator
- XKF, SDRE-KF and EnKF estimates are close

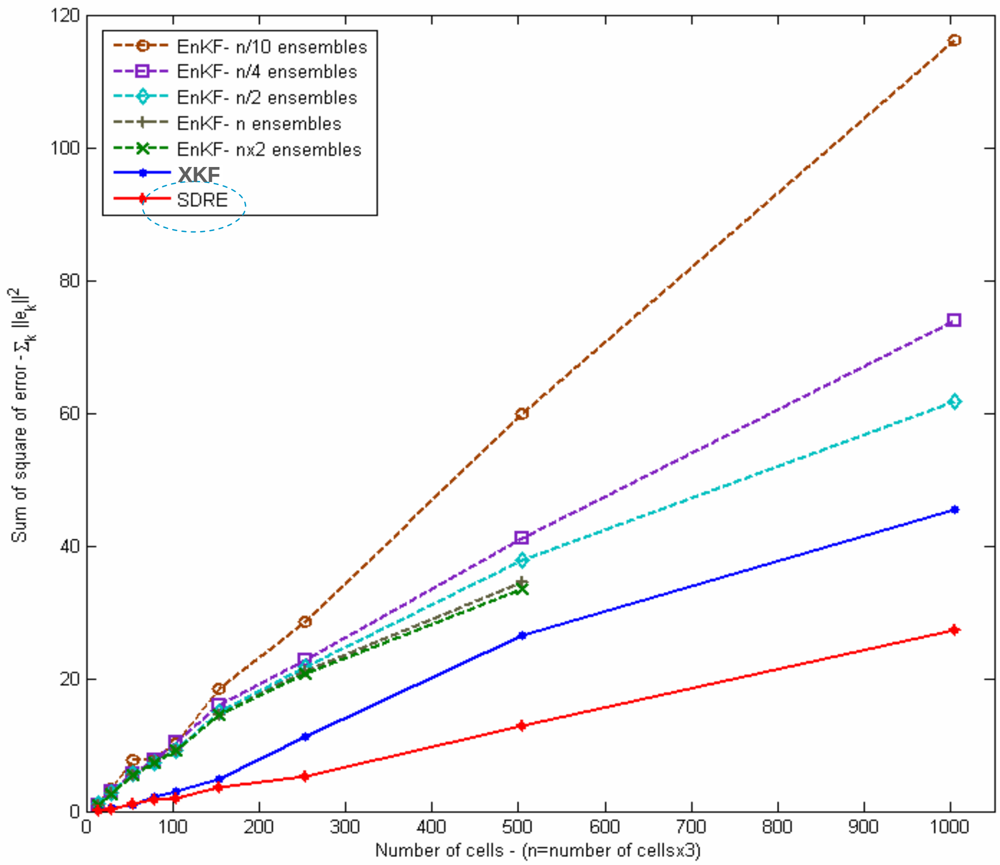
Estimation Performance

- Error in momentum estimates



Estimation Performance

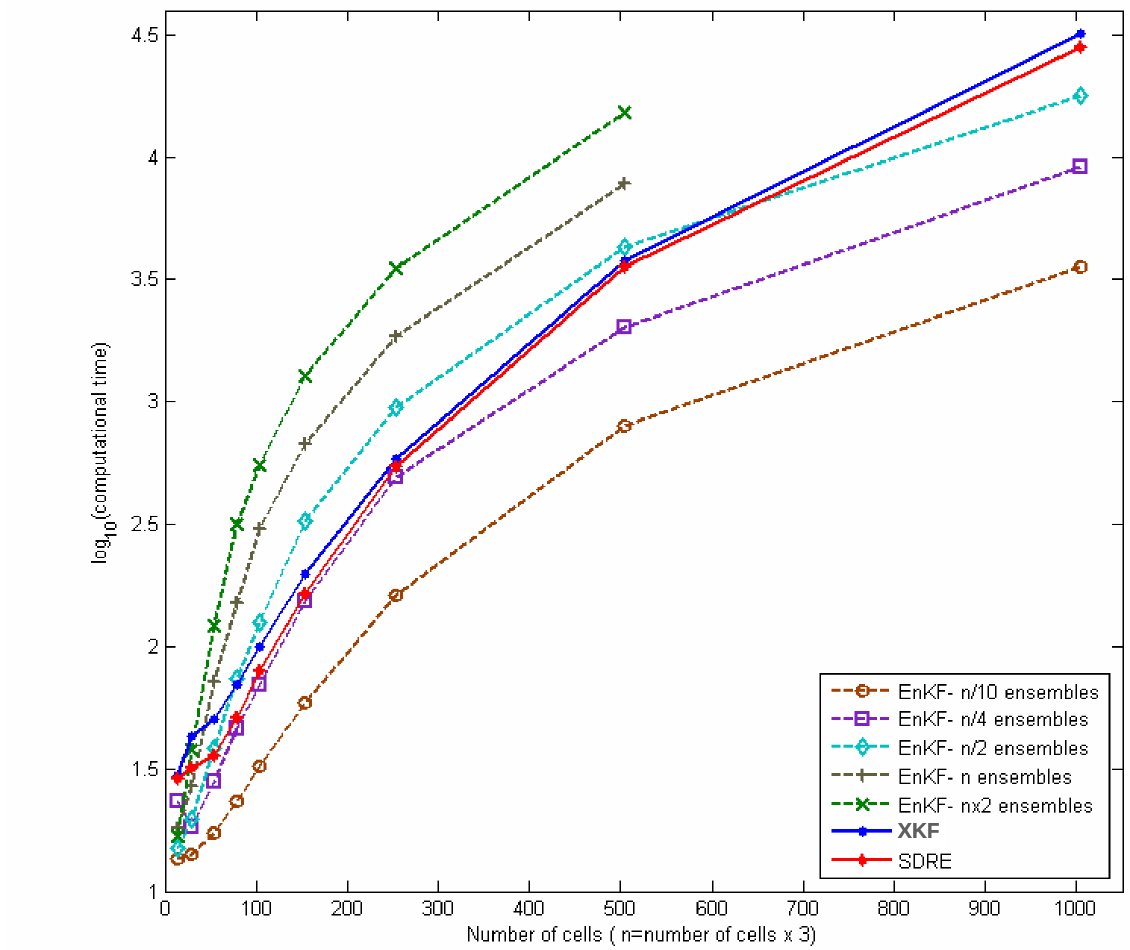
- Size of the grid N varies from 10 cells to 1000 cells
- Ensemble size q of EnKF varies from $N/10$ to $2N$
- Compare the mean-square-error of the state estimates as the grid size increases



Subsonic flow

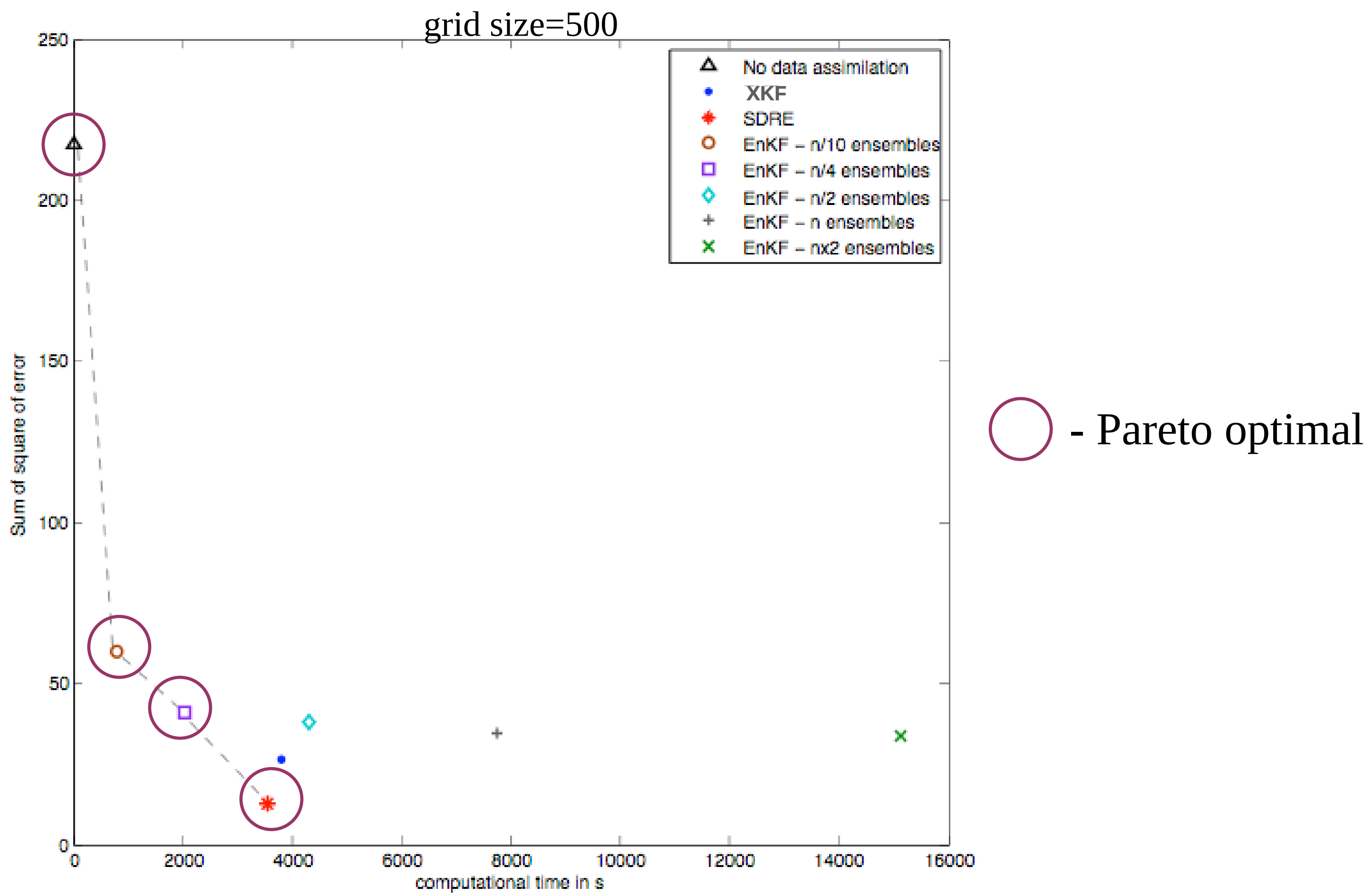
Computational Performance

- Computational time
  - *XKF, SDRE-KF and EnKF*



Computational Performance

- Accuracy *versus* Computational time trade-off



Unknown Dynamics

Model Mismatch

- Assume we use $\hat{A}$ and $\hat{C}$ instead of A and C
- Error dynamics

$$e_{k+1} = (\hat{A} - \widehat{K}_k \hat{C}) e_k + +(\Delta A - \widehat{K}_k \Delta C) x_k + w_k - \widehat{K}_k v_k$$

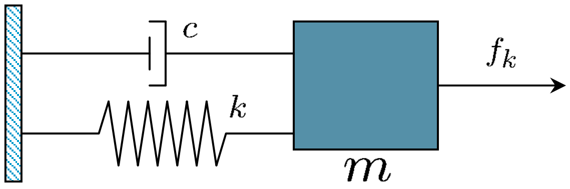
$$\widehat{K}_k = \hat{A} \widehat{P}_k \hat{C} \left(\hat{C} \widehat{P}_k \hat{C}^\top + R \right)^{-1}$$

$$\widehat{P}_{k+1} = \hat{A} \widehat{P}_k \hat{A}^\top - \hat{A} \widehat{P}_k \hat{C}^\top (\hat{C} \widehat{P}_k \hat{C}^\top + R)^{-1} \hat{C} \widehat{P}_k \hat{A}^\top + Q$$

  – The estimates may be biased even if $\hat{x}_0 = \bar{x}_0$
  – $\widehat{P}_k$ is a pseudo error covariance

Estimating Plant Parameters

- Assume certain plant parameters are unknown
  - Mass-spring-damper system
  - Measurement of velocity is available



  - Mass m is unknown

$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} x_{1,k} + t_s x_{2,k} \\ x_{2,k} + t_s(-\frac{k}{m} x_{1,k} - \frac{c}{m} x_{2,k} + f_k) \end{bmatrix}$$

- Kalman filter requires knowledge of system dynamics

- Idea: Estimate position and mass

Estimating Plant Parameters

- Augment the unknown parameter to the state variable
  - $X_k = [x_{1,k} \;\; x_{2,k} \;\; m_k]$

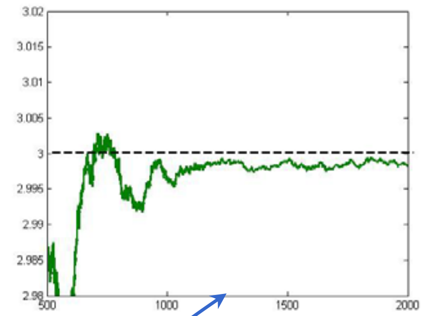
$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \\ m_{k+1} \end{bmatrix} = \begin{bmatrix} x_{1,k} + t_s x_{2,k} \\ x_{2,k} + t_s[-\frac{k}{m_k} x_{1,k} - \frac{c}{m_k} x_{2,k} + f_k] \\ m_k \end{bmatrix}$$

- View the unknown parameter as a state

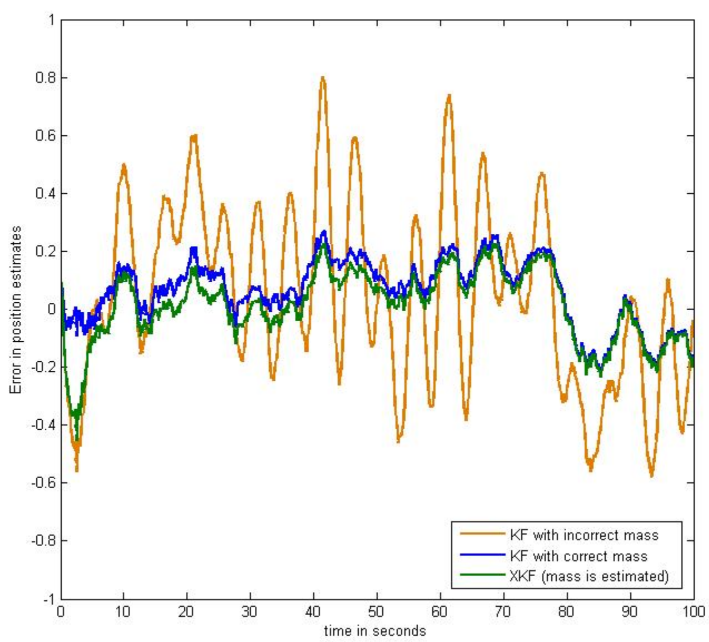
- Use nonlinear estimation techniques to estimate the unknown plant parameter and unmeasured states

Estimating Plant Parameters

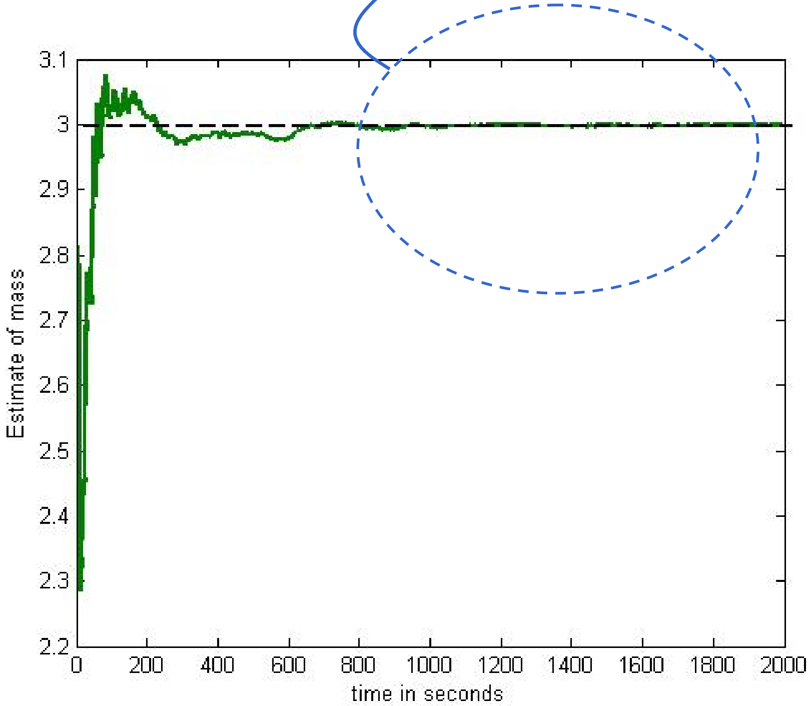
- Compare estimates of position
  - KF with the correct mass
  - KF with incorrect mass estimate
  - XKF (obtain estimates of the mass)



Estimate of mass is inconsistent



Error in position estimates



Estimate of the mass using XKF

Adaptive Estimation

- Asymptotic adaptive observers
  - Noise free conditions
  - Design is easy for single output systems
- Express the system in the observable canonical form
- Estimate the unknown system parameters using direct or indirect methods
- Use the estimates of the parameters in the standard Luenberger observer structure to estimate the state

High Dimension

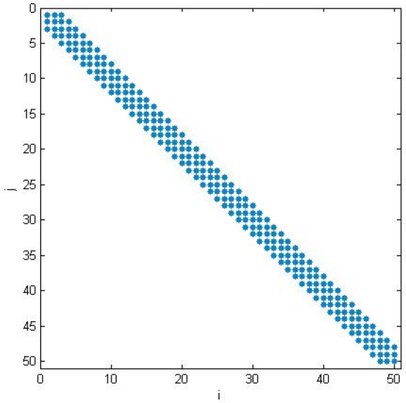
Computational Complexity

- Riccati update of the covariance is computationally expensive
  - $\mathcal{O}(n^3)$ operation
  - n is the dimension of the state variable
  - $n > 1e6$ for weather prediction applications

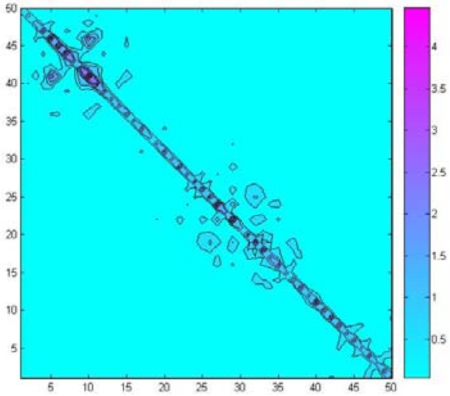
- Techniques for reducing the computational burden
  - Banded covariance
  - Reduced order models
  - Square-root Kalman filtering

Banded Covariance

- Banded dynamics
  - Occurs in systems where the future value of a particular state depends on the current value of only its nearest neighbor states
    - Finite volume discretizations
- Stable linear time-invariant system
  - $P_{k+1} = AP_kA^\mathsf{T} + Q$
  - $\lim_{k\to\infty} P_k = P$ exists



Structure of A
(banded dynamics)



Magnitude of entries of P

Covariance of Banded Dynamics

- Magnitude of the entries of the error covariance progressively decreases as we move away from the diagonal

- The rate of decrement depends on the width of the dynamics (number of nearest neighbors involved)

$$\|H_i \circ P\|_{\mathsf{F}} \leq \frac{\varepsilon^{2i}}{1-\varepsilon^2}\sigma_A \sigma_Q$$

$\|H_i \circ P\|_{\mathsf{F}}$	=	RSS of entries *i* units away from the diagonal
σ_A	=	$\max_i \|A^i\|_{\mathsf{F}}$
σ_Q	=	$\|Q\|_{\mathsf{F}}$
ε	=	< 1

Banded Covariance Approximation

- Neglect correlation between distant cells during data assimilation

- After the Riccati update, retain only the entries of the pseudo error covariance that are within a specified distance from the diagonal

$$\tilde{P}_k = H \circ P_k$$ (Retain only specific entries of the covariance)

$$P_{k+1} = A_k \tilde{P}_k A_k^\top - A_k \tilde{P}_k C_k^\top (C_k \tilde{P}_k C_k^\top + R_k)^{-1} C_k \tilde{P}_k A_k^\top + Q_k$$

$$H_k = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & \cdots \\ 1 & 1 & 1 & 0 & 0 & \cdots \\ 0 & 1 & 1 & 1 & 0 & \cdots \\ & & \ddots & \ddots & & \\ 0 & \cdots & 0 & 1 & 1 & 1 \\ 0 & \cdots & 0 & 0 & 1 & 1 \end{bmatrix}$$ Typical structure of H_k

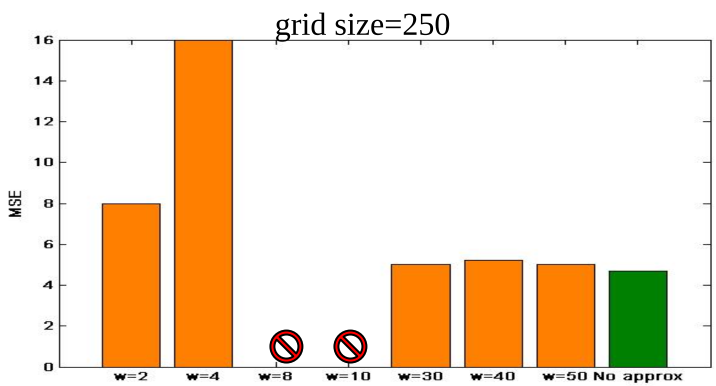
  – Since A_k and $\tilde{P}_k$ are banded diagonal, computational burden of evaluating $A_k \tilde{P}_k A_k^\top$ is reduced

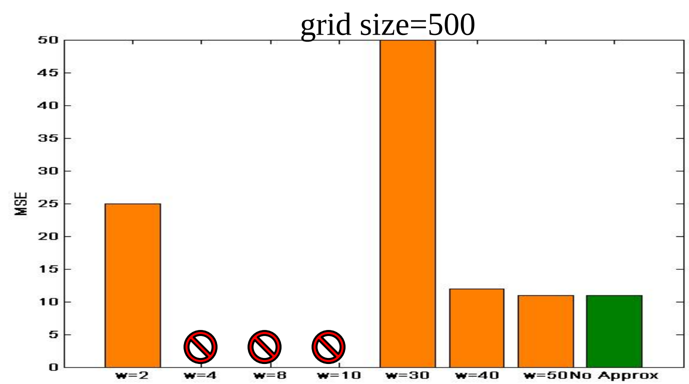
- Positive definiteness of the pseudo error covariance $\tilde{P}_k$ is not guaranteed

  – Retaining large number of entries helps to ensure positive definiteness

Simulation : 1D-Hydrodynamics

- Comparison of error in estimates as the grid size increases
  - Neglect correlation between cells that are farther than distance ω apart
  - Covariance update : $\mathcal{O}(n^3) \longrightarrow \mathcal{O}(\omega^2 n)$





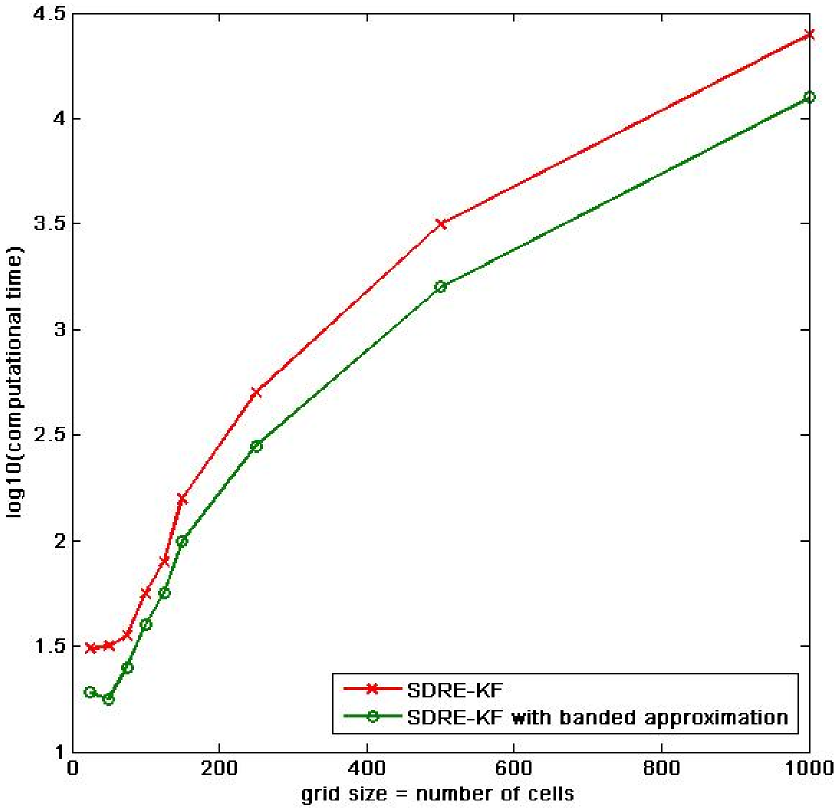
– Error covariance is not positive semidefinite

MSE with SDRE-KF (No approximation)

MSE with SDRE-KF (Banded covariance approximation)

Simulation : 1D- Hydrodynamics

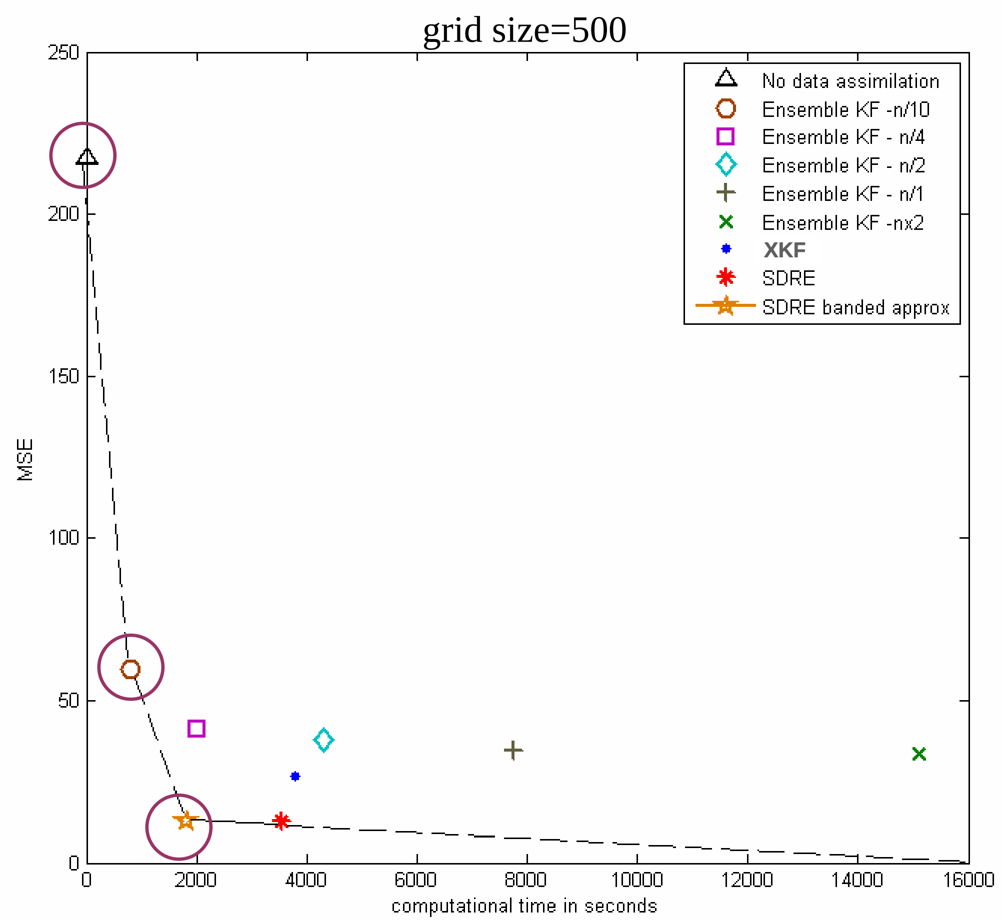
- Compare the time taken for data assimilation



- **Banded covariance approximation reduces the computational time of the SDRE-KF by a factor of 2 (as dimension becomes very large)**
  - No noticeable change in the performance

Simulation : 1D- Hydrodynamics

- Accuracy *versus* Computational time



- Pareto optimal

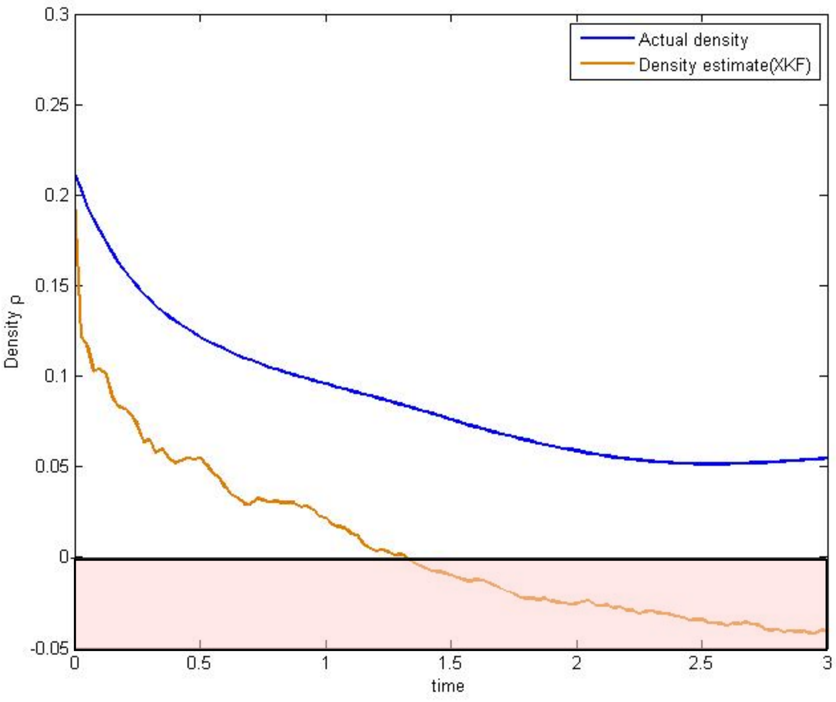
Physical Constraints

State Constraints

- Constraints on states of certain physical systems naturally arise
  - Certain states are always positive
    - Concentration of chemicals
    - Density
    - Kinetic energy
- Do the state estimates also satisfy the same constraints ?

State Constraints

- 1D Hydrodynamics example
  - Density estimates maybe negative !!
  - Results in filter instability



Estimation with Constraints

- Equality constraints

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k + w_k \\ y_k &= C_k x_k + v_k \\ x_{1,k} &= H_k x_k \end{aligned}$$

- View the constraint as a measurement *(Porill,1988)*
  - Estimates the states using the Kalman filter
  - Kalman filter can handle noise-free measurements

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k + w_k \\ \tilde{y}_k &= \begin{bmatrix} C_k \\ H_k \end{bmatrix} x_k + \begin{bmatrix} v_k \\ 0 \end{bmatrix} \end{aligned}$$

Estimation with Constraints

- Inequality constraints
  - Recast estimation as an optimal control problem
  - Use nonlinear programming techniques to solve the optimization problem
- Reduce computational burden by using a moving horizon approach *(Rao, Rawlings, Mayne, 2003)*
  - Ignore old measurements
- Computationally expensive compared to the XKF

Summary

- Kalman filter
  - Provides optimal estimates of the state of a linear time-varying system with stochastic inputs
- Extensions (Open problems)
  - Optimal estimators for nonlinear systems
  - Reducing the computational burden for high dimensional systems
  - Accounting for uncertainty in
    - Noise statistics
    - Dynamics