Systematic Design of Extended Kalman Filters Using Knowledge-Driven Models

Professor Christos Georgakis,

Systems Research Institute and
Department of Chemical and Biological Engineering

Tufts University, Medford, MA 02155

Outline of Talk

- Knowledge-Driven Models
- Open-Loop & Closed-Loop Estimation
- Optimal On-Line Estimation – Kalman Filter
- The Math – A Very Brief Summary
- ***Design Rules – Simple Logic***
- Performance Through Examples

Knowledge-Driven Models

- Instead of "Fundamental" or "First-Principles" Models
- Assumptions on *Inner Workings* of Process
- Material & Energy Balances
  - Momentum Balances, if Needed
- Experimental Data
- Fitting Model to Data
  - Minimize Sum of Squares: $\min_{p_j} \sum_{i=1}^{n} (y_{im} - y_{ie})^2$
  - Estimate Parameters: $p_j = \cdots$
- Finished?

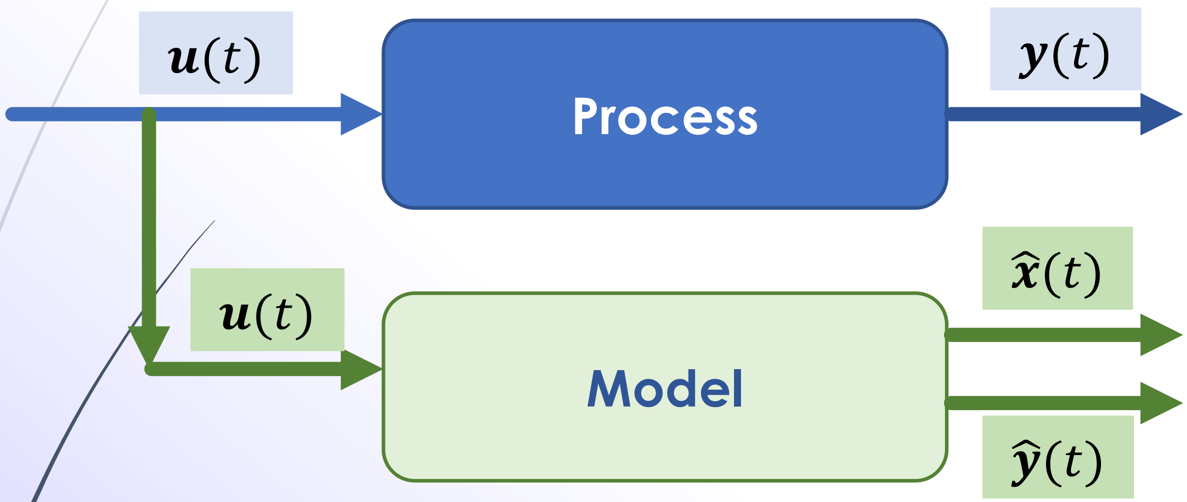
Not Really!

MUST Also Consider

- ARE Parameters Significant?
  - $p_j = 1.2 \pm 0.2$ **NOT** $p_j = 1.2 \pm 2.2$
- Has Model Represented ALL Data?
  - Minimum SS ≅ Normal Variability
    - Minimum SS = SS_{reg}
    - Normal Variability = SS_{pe}
- Lack-of-Fit: $SS_{reg} \cong SS_{pe}$ Is it **OK ?**
- IF NOT: *Revise Model*
- IF YES: *Calculate Covariance Matrix*: C_p

Estimation: Open-Loop

- Use Model and Process Inputs



Is
$y(t) = \hat{y}(t)$
???

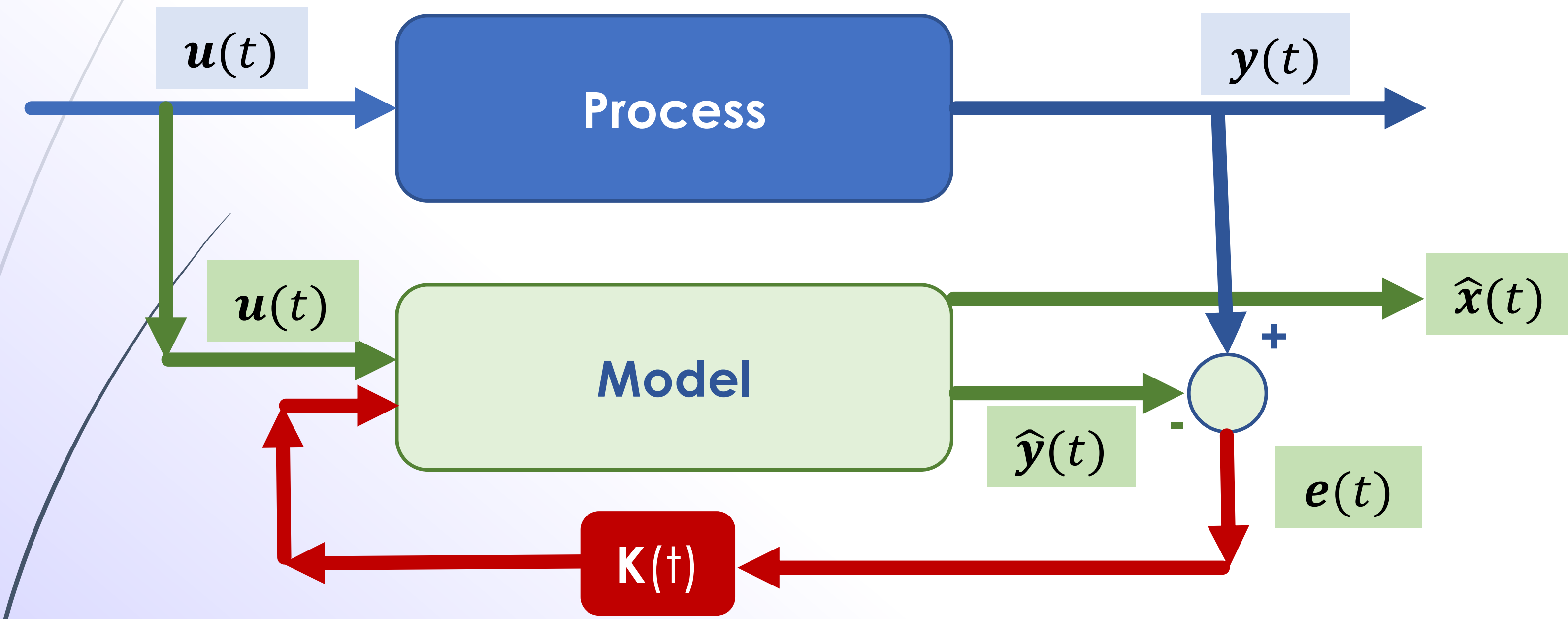
Linear Model: $$\frac{dx}{dt} = \boldsymbol{Ax} + \boldsymbol{Bu};\ \ \boldsymbol{y} = \boldsymbol{Cx}$$

NL Model: $$\frac{dx}{dt} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t);\ \ \boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x})$$

$$\boldsymbol{x} \in \boldsymbol{R}^n, \boldsymbol{y} \in \boldsymbol{R}^m,\ \ m < n$$

Estimation: Closed-Loop

- Use Model and Process Inputs & OUTPUTS



Select K so that $\hat{y}(t) \longrightarrow y(t)$ & $\hat{x}(t) \longrightarrow x(t)$

Design Questions: $\boldsymbol{Q}(t), \boldsymbol{P}(0)$ & $\boldsymbol{R}$

- They Are $n{\times}n$, $n{\times}n$ & $m{\times}m$, Matrices
  - n $x_i(t)$ States --- m $y_j(t)$ Outputs
- Many Authors Think them as Tuning Parameters
  - If n=10, m=3 ➔ 209 parameters
  - If n=50, m=10 ➔5100 parameters
  - … Impossible and Unnecessary Task
- There is only ONE Uncertainty
  - Relative Accuracy of Model vs. Measurements
    - Relative Aggressiveness in Correcting Model
    - Call it ψ .. Needs Tuning

Kalman Filter Math

- Process: $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}\ ) + \boldsymbol{w}(t);\ \boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{v}(t_k)$
  - $\boldsymbol{x}(0) \sim (\widetilde{\boldsymbol{x}}_0, \boldsymbol{P}_0),\ \ \boldsymbol{w}(t_k) \sim (\boldsymbol{0}, \boldsymbol{Q}_k),\ \ \ \boldsymbol{v}(t_k) \sim (\boldsymbol{0}, \boldsymbol{R}_k)$
- Model Parameters:
  - $\boldsymbol{P}_0$: Uncertainty, How Close is $\widetilde{\boldsymbol{x}}_0$ to Real $\boldsymbol{x}(0)$
    - **IF in Doubt**: Use Larger $\boldsymbol{P}_0 = diag\{p_{11,0}, p_{22,0}, \dots, p_{nn,0}\}$
  - $\boldsymbol{Q}_k$: Uncertainty on Model's Accuracy -- **Will Calculate**
  - $\boldsymbol{R}_k$: Accuracy of Measurements -- **We Know**
- Propagation t_k ➔ t_{k+1}^- (Before Measurement Update)
  - $\widehat{\boldsymbol{x}}(t_k^-) = \widehat{\boldsymbol{x}}(t_k^+) + \int_{t_{k-1}}^{t_k} f(\widehat{\boldsymbol{x}}, \boldsymbol{u}, \boldsymbol{p})d\tau\ ;\ \widehat{\boldsymbol{y}}(t_k) = \ \boldsymbol{h}(\widehat{\boldsymbol{x}}(t_k^-))$
- Model Update
  - $\widehat{\boldsymbol{x}}(t_k^+) = \widehat{\boldsymbol{x}}(t_k^-) + \psi \boldsymbol{K}\{\boldsymbol{y}(t_k) - \widehat{\boldsymbol{y}}(t_k)\}$ **…** Usually ψ=1
- But $\boldsymbol{K} = ???$
  - $\boldsymbol{K} = \boldsymbol{0}$ **…** ***IF*** Model is … Perfect

Uncertainty Propagation & $K =$

- $P(t_k^-) = P(t_{k-1}^+) + \int_{t_{k-1}}^{t_k} \{AP + PA^T + Q\}\, d\tau$
  - $A(t) = \partial f/\partial x \quad C(t) = \partial h/\partial x$
- $P(t_k^+) = P(t_k^-) + \delta P(t_k^-)$
  - $\delta P(t_k^-)$=Reduction in Uncertainty =g(**P**, **C**, **K**)
    - *Due to Measurements*
- *Then*
  - $K = P(t_k^-)C^T\{CP(t_k^-)C^T + R\}^{-1}$
    - *Balancing Model $P(t_k^-)$ vs. Measurements R*
    - *Optimal Solution for Linear Systems*
- *WE Are* ***DONE !!!***

The Calculation of $\boldsymbol{Q}_k$

- Model with Parameters $\dot{\hat{\boldsymbol{x}}} = \boldsymbol{f}(\hat{\boldsymbol{x}}, \boldsymbol{u}, \hat{p})$
  - Estimated from Data: $\hat{p}$ and $\boldsymbol{C}_p$ (Covariance)
- Let $J(t) = \partial \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, \widehat{\mathbf{p}}) / \partial \widehat{\mathbf{p}}$
- Then:

$$\boldsymbol{Q}(\mathrm{t}) = J^T(t)\boldsymbol{C}_p J(t)$$

- Ready TO Estimate $\hat{\boldsymbol{x}}(t)$ from Measurement $\boldsymbol{y}(t_k)$
  - **IF System IS Observable**
  - For Linear Systems: rank$[C^T | C^T A^T | .. | C^T (A^T)^{n-1}]$=n

Valappil & Georgakis (2000) AIChE J. 4**6**, p. 292

Is my System Observable?

- Example 1: Small Reaction System
  - 4 Species with 2 Reactions
  - Measure T(t) and Estimate 4 Compositions
    - **Can it Be Done?**
  - YES ***IF Both*** Reactions Affect Energy Balance
- Example 2: Large Reaction System
  - 14 Species with 5 Reactions
  - Measure T(t) and Estimate 14 Compositions
    - **Can it Be Done?**
  - *In Principle* YES … ***IF ALL*** Rxs Affect Energy Balance

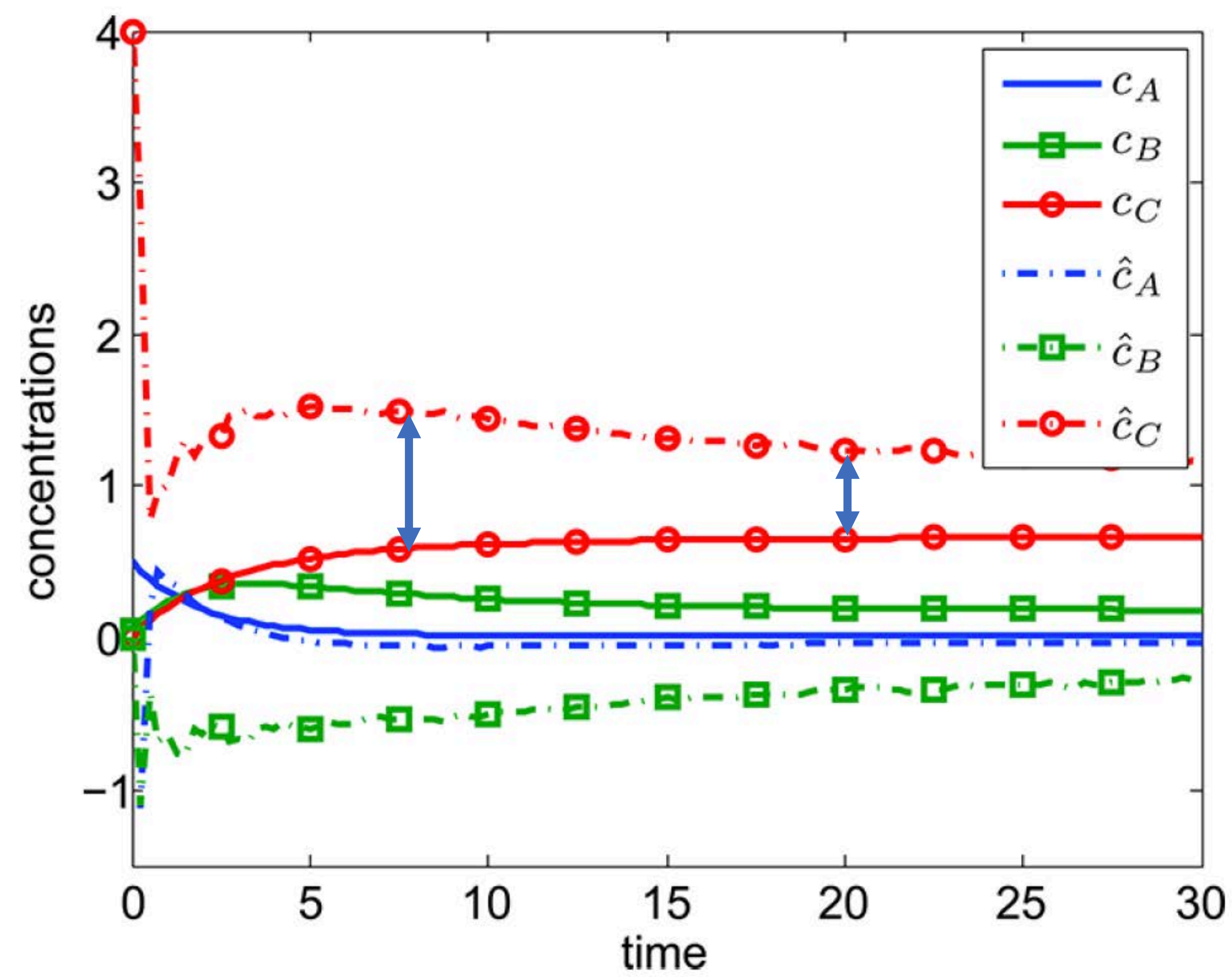
→ We are asking **TOO** Much

Batch Reactor - Example

- Two Reactions (H&R: Haseltine & Rawlings 2005)
  - $A \rightleftarrows B + C \quad r_1 = k_1 A - k_2 BC$ -- $k_1 = 0.5 \quad k_2 = 0.05$
  - $2B \rightleftarrows C \quad r_2 = k_3 B^2 - k_4 C$ -- $k_3 = 0.2 \quad k_4 = 0.01$
- Model: $(x_1 = A,\ x_2 = B,\ x_3 = C)$
  - $dx_1/dt = -r_1; \quad dx_3/dt = r_1 - r_2; \quad dx_3/dt = r_1 + r_2$
  - Measurement: $y = p_t \quad y = \alpha(x_1 + x_2 + x_3)$
- Process Starts at t=0: $(x_1, x_2, x_3) = (0.5,\ 0.05, 0.0)$
- ***H&R*** Start EKF at:
  - $(\hat{x}_1, \hat{x}_2, \hat{x}_3) = (0, 0, 4)$!!! & $P_0 = diag(0.5^2, 0.5^2, 0.5^2)$
    - $\hat{x}_1 = 0 \pm 1.01$ but $x_1 = 0.5$ → **OK**
    - $\hat{x}_2 = 0 \pm 1.01$ but $x_2 = 0.05$ → **OK**
    - $\hat{x}_3 = 4 \pm 1.01$ but $x_1 = 0$ → **NOT OK**

Filter Fails !!!

EKF Fails: $\hat{C}(\infty) \neq C(\infty)$

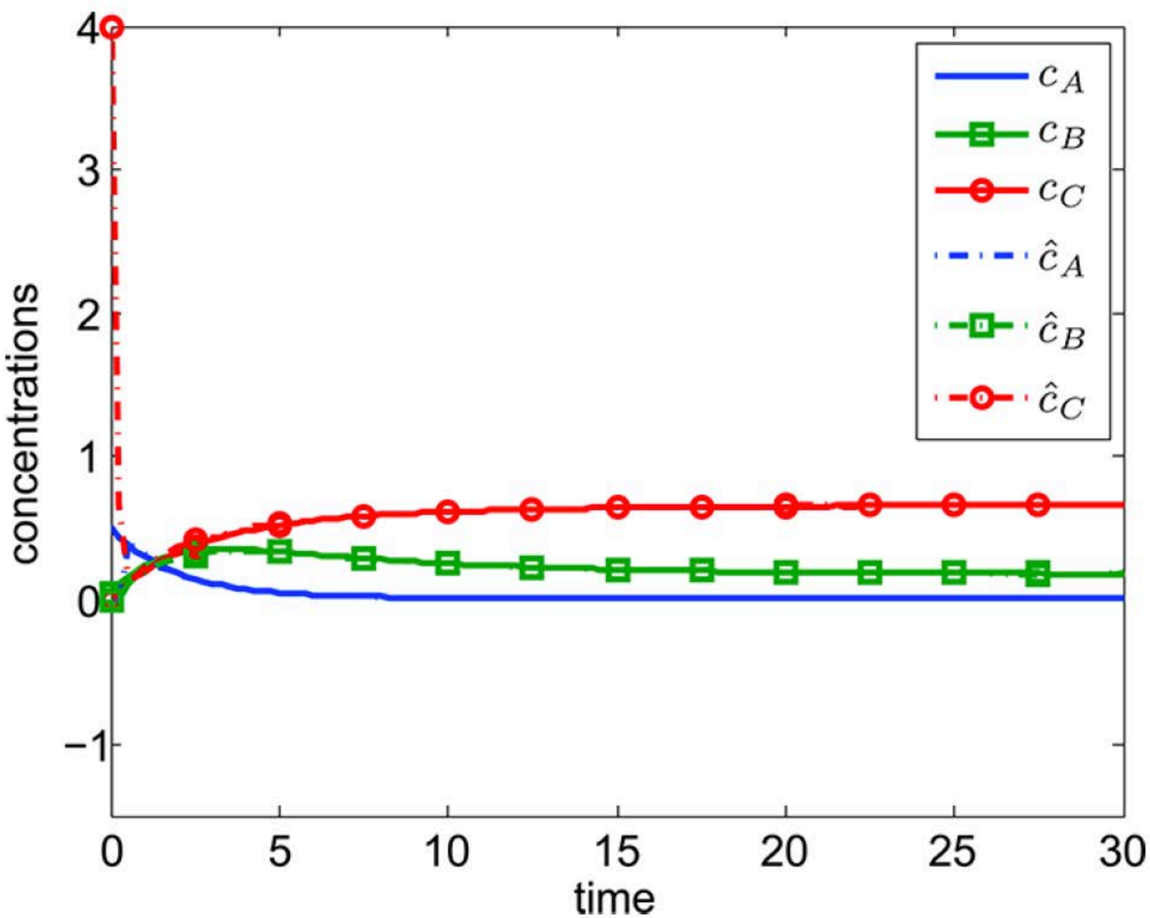


- Haseltine & Rawlings (2005): *Ind. & Eng. Chem. Res.* ***44***, p. 2451
- Schneider & Georgakis (2013): "How To NOT Make the Extended Kalman Filter Fail" *Ind. & Eng. Chem. Res.* **52**, p. 3354

Select $\boldsymbol{P}(0)$ Correctly

True Values: $(x_1, x_2, x_3) = (0.5,\ 0.05, 0.0)$

- From: $(\hat{x}_1, \hat{x}_2, \hat{x}_3) = (0, 0, 4)$!!! & $P_0 = diag(0.5^2, 0.5^2, 0.5^2)$
- TO: $(\hat{x}_1, \hat{x}_2, \hat{x}_3) = (0, 0, 4)$!!! & $P_0 = diag(0.5^2, 0.05^2, 4^2)$
  - $\hat{x}_3 = 4 \pm 16.04$ but $x_1 = 0$ → **NOW OK**



1000 Simulations with 2 $\boldsymbol{P}(0)$s

Table 1. Number of Converging Runs and Estimation Error Statistics[a]

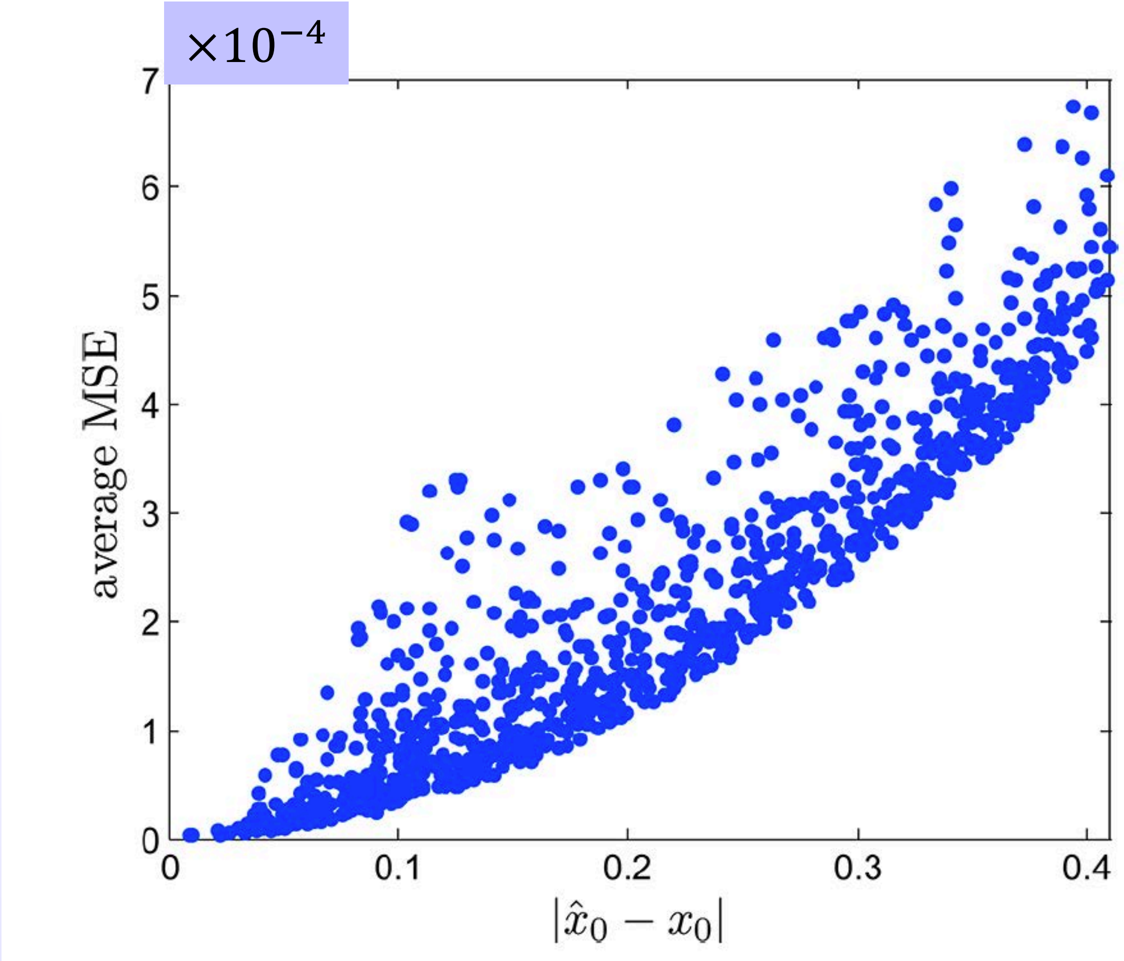
$diag(\boldsymbol{P}_0)$			conv. runs	avg MSE	min MSE	max MSE	std MSE
0.5^2	0.5^2	0.5^2	205	0.3331	0.0803	0.4331	0.1218
0.5^2	0.05^2	4^2	1000	0.0468	0.0465	0.0478	0.0002

[a] For each choice of the diagonal $\boldsymbol{P}_0$ matrix, 1000 EKF runs with different measurement noise realizations were performed.

The $\boldsymbol{P}(0)$ Value is Important

If in Doubt, Select a Larger Value.

Estimation Error vs. Initial Error



Same Can Be Said about the $\boldsymbol{Q}$ Matrix

KFs or EKFs: What Holds You Back?

- Do You have a Knowledge-Driven Model ?
  - *If YES, Use a KF or EKF*
  - If NOT, then What?
- Develop a Data-Driven Model
  - Using ML or "Deep" Learning
- Use Design of Dynamic Experiments (DoDE)
  - Georgakis (2013) I&ECR.
- Use Dynamic Response Methodology (DRSM)
  - Klebanov & Georgakis (2016) I&ECR.
  - More Publications with Pfizer and Merck Colleagues.

Thank You Very Much for Your Attention