Nonlinear Bayesian State Estimation: A Review of Recent Developments

Sachin C. Patawardhan
Department of Chemical Engineering
I.I.T. Bombay
Email: sachinp@iitb.ac.in

Outline

- Motivation and Origin
- Nonlinear State Estimation
  - Extended Kalman Filter
  - Deterministic Derivative-free estimators
  - Particle Filters
- Constrained State Estimation
- Estimation under Model-Plant Mismatch
  - On-line Model Maintenance
- Future research directions

Notation

Mechanistic Model

State Dynamics

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x},\mathbf{u},\mathbf{d},\theta)$$

Measurement Model

$$\mathbf{y} = H[\mathbf{x}]$$

Assumptions

Manipulated inputs and piecewise constant

$$\mathbf{u}(\mathrm{t}) = \mathbf{u}(k) \quad \text{for} \quad t_{\mathrm{k}} \le t < t_{k+1} = t_k + T$$

Unmeasured disturbances are modelled as piecewise constat random fluctuations in the neighborhood of mean value

$$\mathbf{d}(t) = \overline{\mathbf{d}} + \mathbf{w}(k) \quad \text{for} \quad t_{\mathrm{k}} \le t < t_{k+1}$$

Notation

$$\mathbf{x}(t_{k+1}) = \mathbf{x}(t_k) + \int_{t_k}^{t_{k+1}} \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(k), \bar{\mathbf{d}} + \mathbf{w}(k), \boldsymbol{\theta})\, d\tau$$

$$t_k = kT \qquad t_{k+1} = (k+1)T \qquad T : Sampling\ Time$$

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \int_{t_k}^{t_{k+1}} \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(k), \bar{\mathbf{d}} + \mathbf{w}(k), \theta)\, d\tau$$

$$= F\left[\mathbf{x}(k), \mathbf{u}(k), \mathbf{w}(k), \boldsymbol{\theta}\right]$$

Control Relevant Discrete Time Representation

$$\mathbf{x}(k+1) = F\left[\mathbf{x}(k), \mathbf{u}(k), \mathbf{w}(k), \boldsymbol{\theta}\right]$$

$$\mathbf{y}(k) = H\left[\mathbf{x}(k)\right]$$

Bayesian Formulation

Automation Lab
IIT Bombay

Models: mechanistic models of the form

$$\mathbf{x}(k+1) = F\left[\mathbf{x}(k), \mathbf{u}(k), \mathbf{w}(k), \boldsymbol{\theta}\right]$$

$$\mathbf{y}(k) = H\left[\mathbf{x}(k)\right] + \mathbf{v}(k)$$

w(k): uncertainty in states due to unknown inputs

v(k): measurement errors (noise)

(stationary random processes with known statistical properties)

Objective

Find the conditional probability density function (PDF),

$$p\left[\mathbf{x}(k) | \mathbf{Y}^k\right]$$

Y^k :set of all the available measurements up to time instant k.

Nonlinear Bayesian Estimation

Alternative Approaches

- **Sequential Unconstrained Estimation:** Methods that obtain the conditional density function by application of Bayes' rule, and then obtain the estimate using one of the optimization criteria
- **Direct Optimization:** Methods that assume a suitable form for the prior probability density function and convert the estimation problem directly into an optimization problem.
  - Sequential constrained estimators
  - Moving horizon estimator

Sequential Bayesian Estimation

Automation Lab
IIT Bombay

Prediction step: posterior density at previous time step is propagated into next time step through state transition density to compute prior

$$p\left[\mathbf{x}(k) \mid \mathbf{Y}^{k-1}\right] = \int p\left[\mathbf{x}(k) \mid \mathbf{x}(k-1)\right] p\left[\mathbf{x}(k-1) \mid \mathbf{Y}^{k-1}\right] d\mathbf{x}(k-1)$$

Update step: Computation of posterior density from the prior

$$p\left[\mathbf{x}(k) \mid \mathbf{Y}^{k}\right] = \frac{p\left[\mathbf{y}(k) \mid \mathbf{x}(k)\right]}{p\left[\mathbf{y}(k) \mid \mathbf{Y}^{k-1}\right]} \times p\left[\mathbf{x}(k) \mid \mathbf{Y}^{k-1}\right]$$

The Posterior Density function constitutes the complete solution to the sequential estimation problem.

Bayesian Estimation

- Prediction and update strategy provides an optimal solution to the state estimation problem
  - involves high-dimensional integration.
  - exact analytical solution to the recursive propagation of the posterior density is difficult obtain
- Linear state estimation: possible to compute analytical solution
- Nonlinear filtering techniques: develop **approximate** and **computationally tractable sub-optimal (local) solutions** to the sequential Bayesian estimation problem

Approximation Approaches

- Prediction step
  - Taylor series approximation
  - Deterministic sampling based approximations
  - Stochastic sampling (Monte Carlo) based approximations

- Update step
  - Statistical linear regression or linear minimum mean square estimation
  - Monte Carlo sampling based approximations

Extended Kalman Filter (EKF)

- Most popular and widely used Nonlinear Bayesian Filter
- Propagation step: Predicted Mean

$$
\begin{aligned}
\hat{\mathbf{x}}(k \mid k-1) &= E\left[\mathbf{x}(k) \mid Y^{k-1}\right] \\
&= E\left[F\left[\mathbf{x}(k-1), \mathbf{u}(k-1), \mathbf{w}(k-1)\right] \mid Y^{k-1}\right]
\end{aligned}
$$

Using Taylor series approximation in the nbhd of $(\bullet) \equiv \left[\hat{\mathbf{x}}(k-1 \mid k-1), \mathbf{u}(k-1), \bar{0}\right]$

$$
\begin{aligned}
F\left[\mathbf{x}(k-1), \mathbf{u}(k-1), \mathbf{w}(k-1)\right] &\approx F\left[\hat{\mathbf{x}}(k-1 \mid k-1), \mathbf{u}(k-1), \bar{0}\right] \\
&+ \left[\frac{\partial F}{\partial \mathbf{x}}\right]_{(\bullet)} \boldsymbol{\varepsilon}(k-1 \mid k-1) + \left[\frac{\partial F}{\partial \mathbf{d}}\right]_{(\bullet)} \mathbf{w}(k-1)
\end{aligned}
$$

$$\Downarrow$$

$$
\begin{aligned}
E\left[F\left[\mathbf{x}(k-1), \mathbf{u}(k-1), \mathbf{w}(k-1)\right] \mid Y^{k-1}\right] &\approx F\left[\hat{\mathbf{x}}(k-1 \mid k-1), \mathbf{u}(k-1), \bar{0}\right] \\
&= F\left[E\left[\mathbf{x}(k-1) \mid Y^{k-1}\right], \mathbf{u}(k-1), \bar{0}\right]
\end{aligned}
$$

Extended Kalman Filter (EKF)

Predicted Covariance

$$\mathbf{P}(k|k-1) \approx \left[\frac{\partial F}{\partial \mathbf{x}}\right]_{(\bullet)} \mathbf{P}(k-1|k-1)\left[\frac{\partial F}{\partial \mathbf{x}}\right]_{(\bullet)}^{T} + \left[\frac{\partial F}{\partial \mathbf{d}}\right]_{(\bullet)} \mathbf{Q}\left[\frac{\partial F}{\partial \mathbf{d}}\right]_{(\bullet)}^{T}$$

- Update Step: Updated Mean computation using Statistical Linear Regression

$$\mathbf{L}(k) = \mathbf{P}_{\varepsilon\mathbf{e}}(k)\left[\mathbf{P}_{\mathbf{ee}}(k)\right]^{-1}$$

$$\mathbf{e}(k) = \mathbf{y}(k) - H\left[\hat{\mathbf{x}}(k|k-1)\right]$$

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \mathbf{L}(k)\mathbf{e}(k)$$

$$\mathbf{P}_{ee}(k) \approx \left[\frac{\partial H}{\partial \mathbf{x}}\right]_{(\bullet)} \mathbf{P}(k|k-1)\left[\frac{\partial H}{\partial \mathbf{x}}\right]_{(\bullet)}^{T} + \mathbf{R}$$

$$\mathbf{P}_{\varepsilon e}(k) \approx \mathbf{P}(k|k-1)\left[\frac{\partial H}{\partial \mathbf{x}}\right]_{(\bullet)}^{T}$$

EKF: Update Step

Updated Covariance

$$\mathbf{P}(k|k) = \left(\mathbf{I} - \mathbf{L}(k) \left[\frac{\partial H}{\partial \mathbf{x}} \right]_{(\bullet)} \right) \mathbf{P}(k|k-1)$$

- Approximates p[x(k)|Y^k] and p[x(k)|Y^{k-1}] to be Gaussian i.e.

$$p\left[\mathbf{x}(k) | \mathbf{Y}^{k-1}\right] \approx \mathrm{N}(\hat{\mathbf{x}}(k|k-1), \mathbf{P}(k|k-1))$$

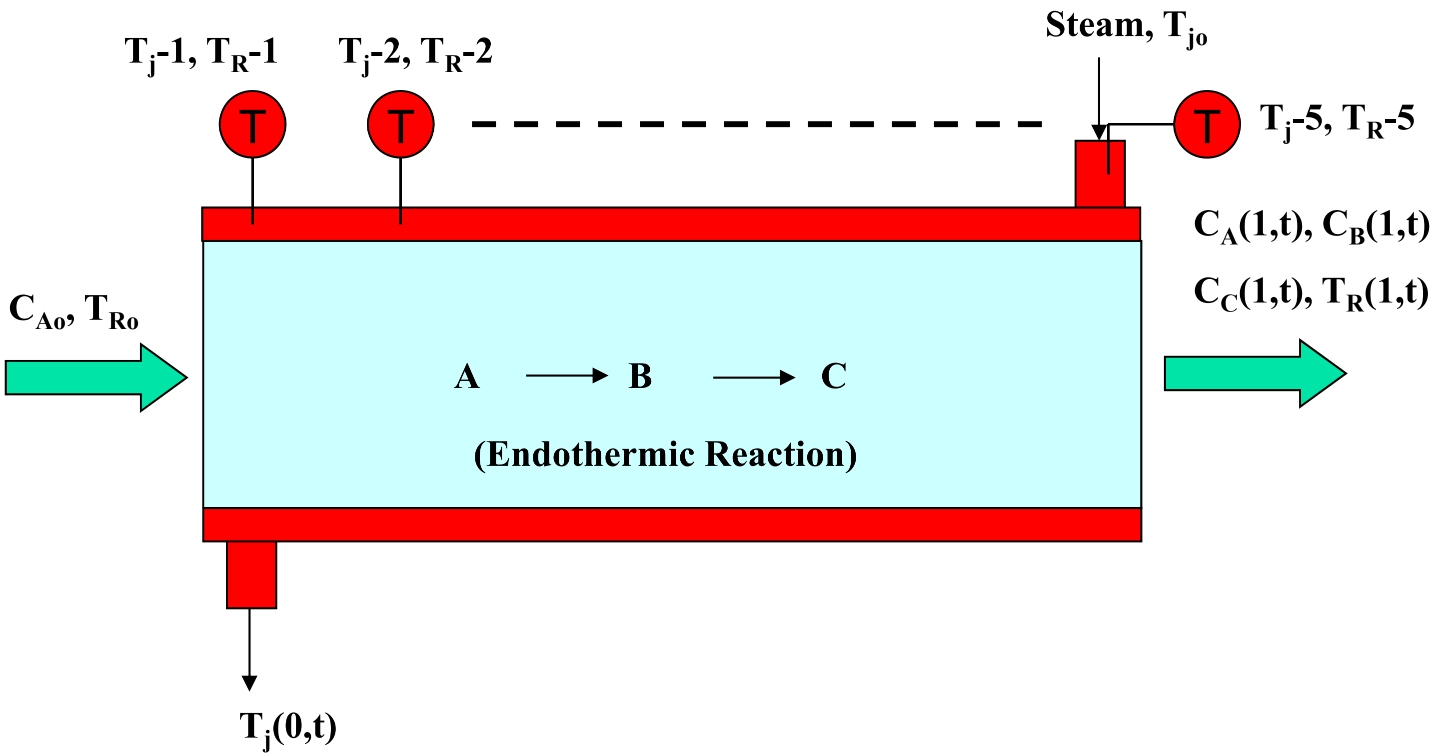
and

$$p\left[\mathbf{x}(k) | \mathbf{Y}^{k}\right] \approx \mathrm{N}(\hat{\mathbf{x}}(k|k), \mathbf{P}(k|k))$$

Gaussian approximation: simplest method to approximate numerical integration problem due to its analytical tractability

Local asymptotic convergence of estimation error (in absence of the state and the measurement noise) has been established using Lyapunov's second method (Reif at al. 1999)

EKF : Plug Flow (Tubular) Reactor (PFR)



State Estimation Problem
Estimate concentration profile inside the reactor using few temperature measurements along the length

Fixed Bed Reactor

- **Material Balances** **(Distributed Parameter System)**

$$\frac{\partial C_A}{\partial t} = -v_l \frac{\partial C_A}{\partial z} - k_{10} e^{-E_1/RT_r} C_A$$

........Reactant A

$$\frac{\partial C_B}{\partial t} = -v_l \frac{\partial C_B}{\partial z} + k_{10} e^{-E_1/RT_r} C_A - k_{20} e^{-E_2/RT_r} C_B$$

........Product B

- **Energy Balances**

$$\frac{\partial T_r}{\partial t} = -v_l \frac{\partial T_r}{\partial z} + \frac{(-\Delta H_{r1})}{\rho_m C_{pm}} k_{10} e^{-E_1/RT_r} C_A$$

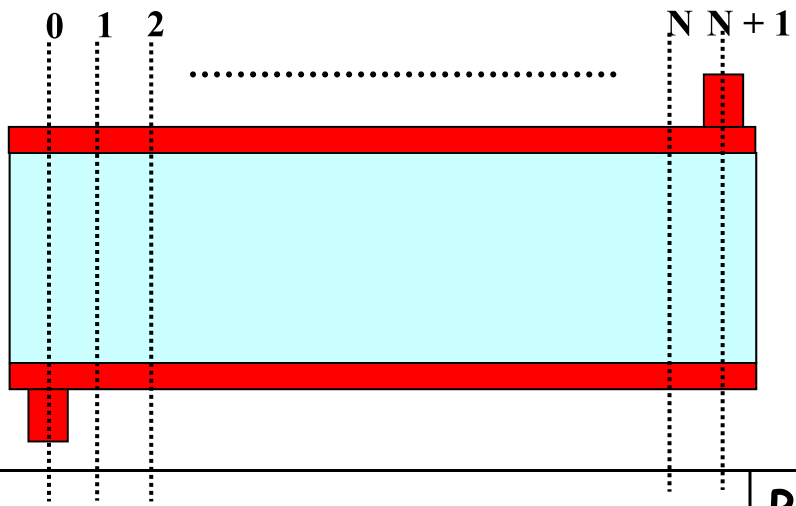
$$+ \frac{(-\Delta H_{r2})}{\rho_m C_{pm}} k_{20} e^{-E_2/RT_r} C_B + \frac{U_w}{\rho_m C_{pm} V_r} \left(T_j - T_r\right)$$

........Reactor Temp.

$$\frac{\partial T_j}{\partial t} = u \frac{\partial T_j}{\partial z} + \frac{U_{wj}}{\rho_{mj} C_{pmj} V_j} \left(T_r - T_j\right)$$

........Jacket Temp.

PDE To ODE Model (Finite Differencing)



	Plant	Model
No. of internal discretization points	19	4
No. of states	80	20
No. of jacket side temp. measurements	3	3
No. of reactor side temp. measurements	3	3

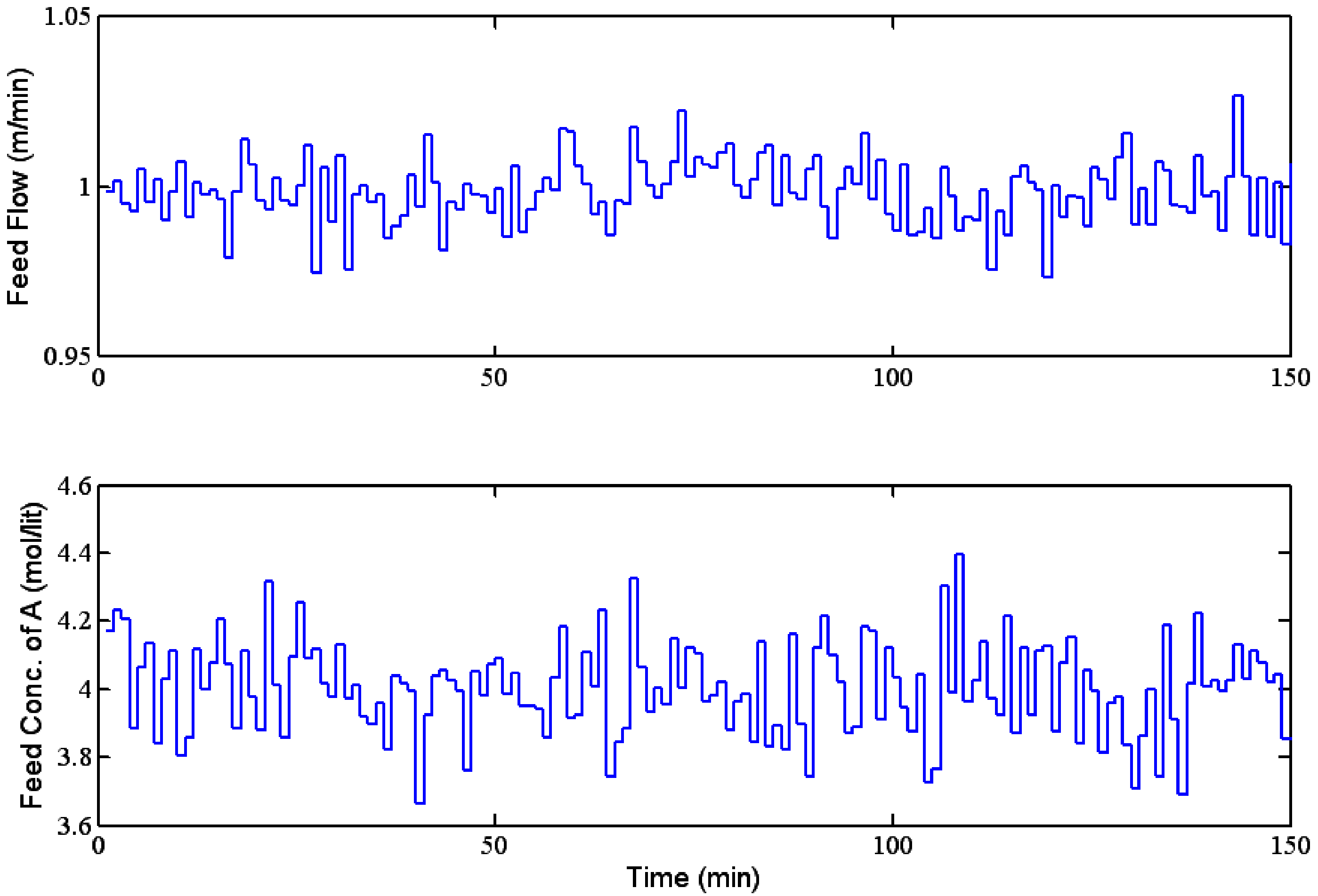
State Estimation using EKF

Simulation Parameters

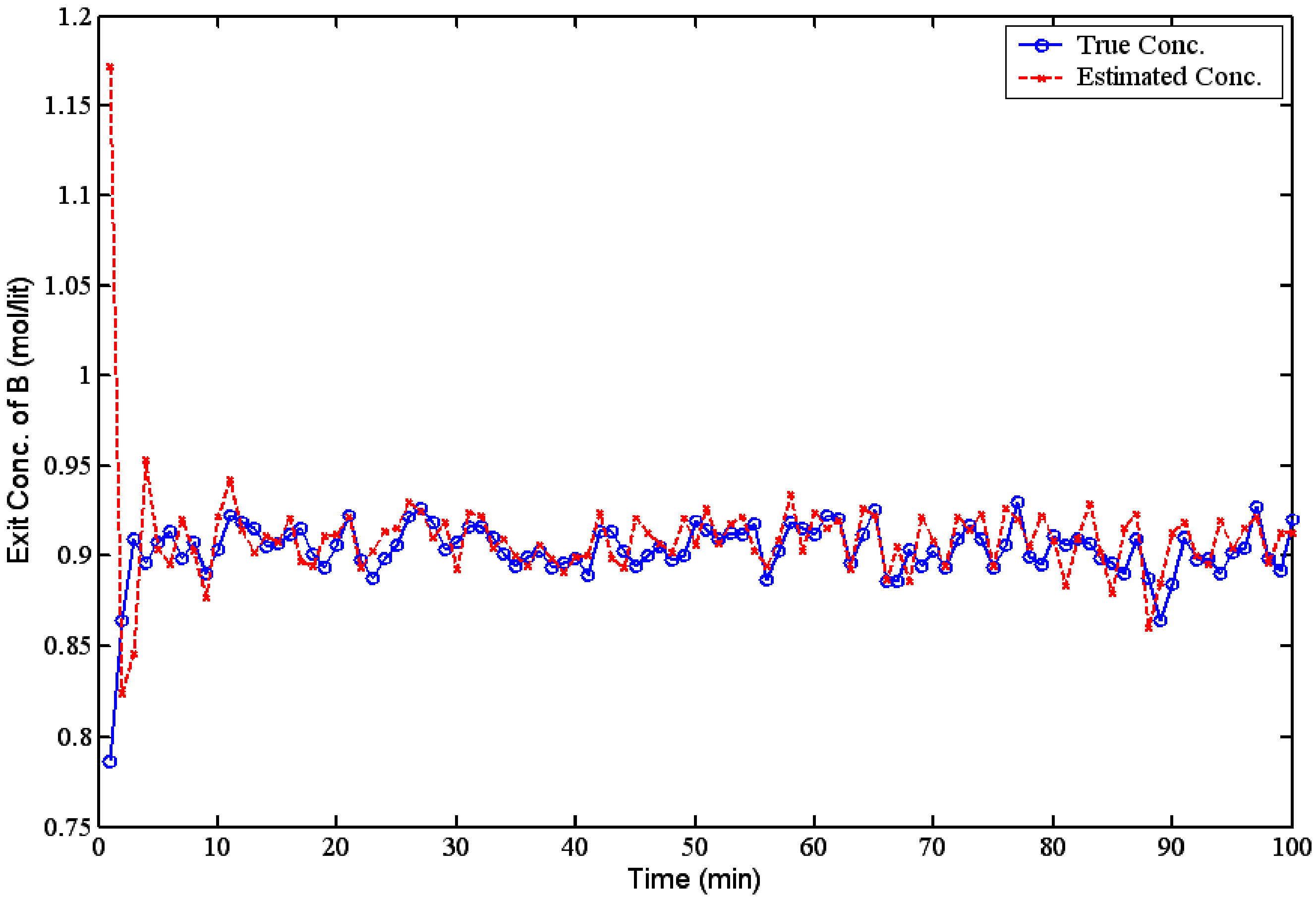
Variable	Nominal Value	Fluctuations added
Feed Flow	1 m/min	0.01 m/min
Feed Concentration	4 mol/lit	0.14 mol/lit
Temperature measurements	-	0.4 K
Steam flow rate	1 m/min	-

- Performance of EKF under the effect of feed flow and feed concentration fluctuations was studied
- The estimated concentration approaches the true concentration within 5 minutes

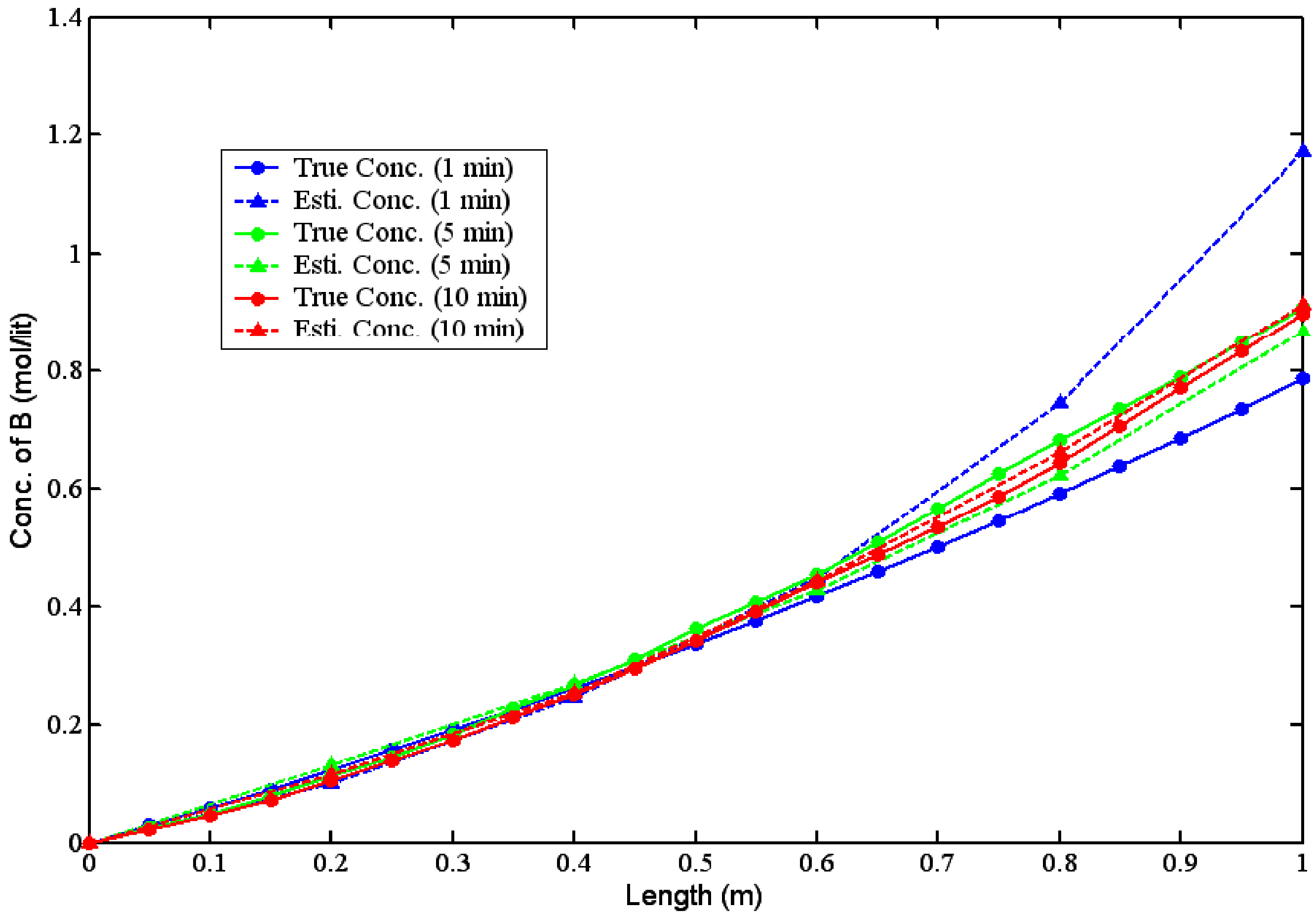
Fluctuations in Feed Flow and Feed Concentration



Actual and Estimated Exit Concentration of B



Simulation Result: Concentration profiles of product B at different time instants



State and Parameter Estimation

Estimation of deterministic changes in unmeasured disturbances / model parameters

$$\mathbf{X}(k+1) = \mathbf{X}(k) + \int_{kT}^{(k+1)T} \mathbf{F}\left[\mathbf{X}(\tau), \mathbf{U}(k), \boldsymbol{\theta}(k)\right] d\tau + \mathbf{w}(k)$$

$$\boldsymbol{\theta}(k+1) = \boldsymbol{\theta}(k) + \mathbf{w}_{\boldsymbol{\theta}}(k)$$

$$\mathbf{Y}(k) = \mathbf{H}\left[X(k)\right] + \mathbf{v}(k)$$

Augment the model with fictitious discrete evolution equation

$\boldsymbol{\theta}(k)$: Vector containing unmeasured disturbances / parameters to be estimated with states

State and Parameter Estimation

Prediction step:

$$\begin{bmatrix} \hat{\mathbf{X}}(k|k-1) \\ \hat{\boldsymbol{\theta}}(k|k-1) \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{X}}(k-1|k-1) + \int_0^T \mathbf{F}\left[\mathbf{X}(\tau), \mathbf{U}(k-1), \boldsymbol{\theta}(k-1|k-1)\right]dt \\ \hat{\boldsymbol{\theta}}(k-1|k-1) \end{bmatrix}$$

Correction Step:

$$\begin{bmatrix} \hat{\mathbf{x}}(k|k) \\ \hat{\boldsymbol{\theta}}(k|k) \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{x}}(k-1|k-1) \\ \hat{\boldsymbol{\theta}}(k-1|k-1) \end{bmatrix} + \mathbf{L}(k)\left[\mathbf{y}(k) - \mathbf{C}\hat{\mathbf{x}}(k|k-1)\right]$$

Covariance Update

$$\mathbf{A}(k) = \left[\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right]_{(\bullet)} \quad ; \quad \mathbf{B}_{\boldsymbol{\theta}}(k) = \left[\frac{\partial \mathbf{F}}{\partial \boldsymbol{\theta}}\right]_{(\bullet)}$$

$$(\bullet) \equiv (\mathbf{X}(k-1|k-1), \mathbf{U}(k-1), \boldsymbol{\theta}(k-1|k-1))$$

State and Parameter Estimation

Covariance Update : using augmented matrices

$$\Phi(k) = \exp[TA(k)] \; ; \; \Gamma_\theta(k) = \int_0^T \exp[A(k)\tau] B_\theta(k) d\tau$$

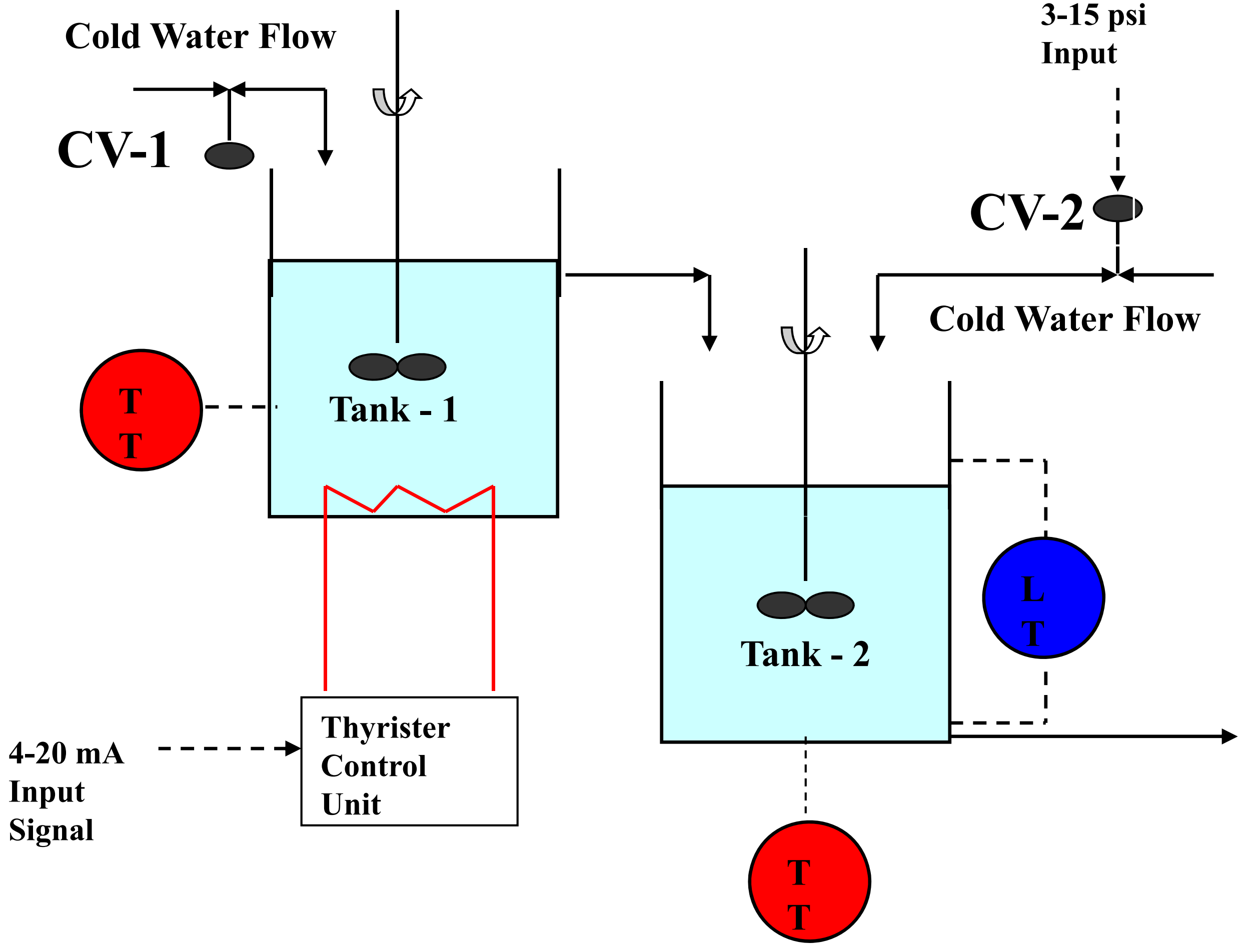
$$\Phi_a(k) = \begin{bmatrix} \Phi(k) & \Gamma_\theta(k) \\ [0] & [0] \end{bmatrix}$$

$$\text{State Noise Covarance :} \begin{bmatrix} Q & [0] \\ [0] & Q_\theta \end{bmatrix}$$

Tuning Parameter (Q_θ)

Fast changing parameter / disturbance : use high values of co-variance

Experiment: Combined State and Parameter Estimation on Heater-Mixer Setup



Example: Stirred Tank Heater-Mixer

$$\frac{dT_1}{dt} = \frac{F_1}{V_1}(T_{i1} - T_1) + \frac{\beta Q(I_1)}{V_1 \rho C_p}$$

Parameter to be estimated simultaneously with states

β : Heat - loss factor

$$\frac{dh_2}{dt} = \frac{1}{A_2}\left[F_1 + F_2(I_2) - F\right]$$

$$\frac{dT_2}{dt} = \frac{1}{h_2 A_2}\left[F_1(T_1 - T_2) + F_2(T_{i2} - T_2) - \frac{UA(T_2 - T_{atm})}{\rho C_p}\right]$$

$$Q(I_1) = 7.979 I_1 + 0.989 I_1^2 - 0.0073 I_1^3$$

$$F_2(I_2) = 3.9 + 27 I_2 - 0.71 I_2^2 + 0.0093 I_2^3$$

$$U = 139.5\, J / m^2 {}^0 Ks \quad ; \quad F(h) = k\sqrt{h_2 - \bar{h}}$$

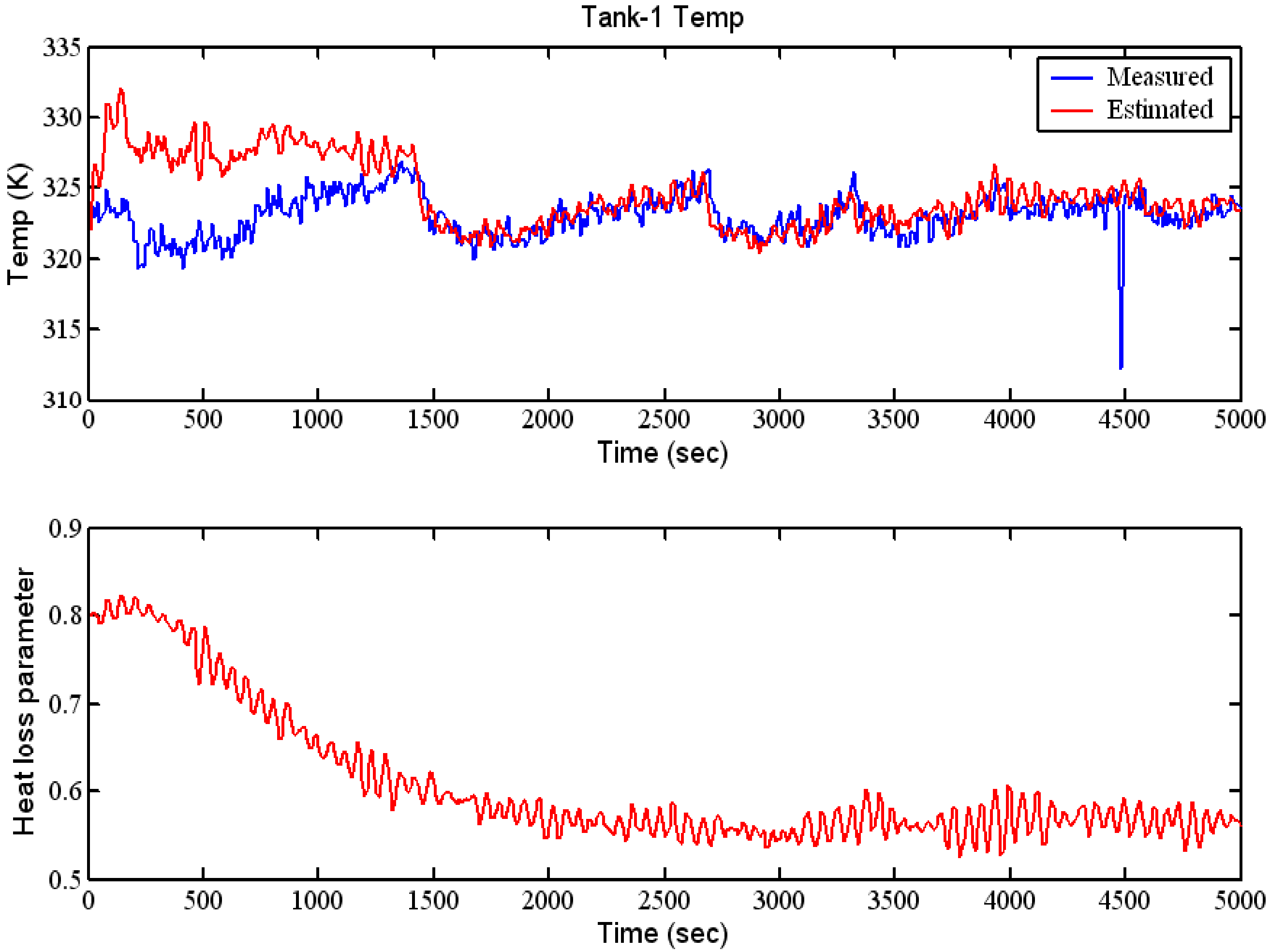
I_1 : % current input to thyrister power controller

I_2 : % current input to control valve

Estimation of states and parameters using EKF

- Tank 1 temperature and heat loss parameter are to be estimated using EKF
- Tank 2 temperature and level are measured
- The system is kept in perturbed state by perturbing the inputs (heater input and tank 2 inlet flow)
- The flow to tank 1 is kept constant.
- The heat-loss parameter (β)is initialized with a value of 0.8

Experimental result: Tank 1 temperature and heat loss parameter estimates



Outline

- Motivation and Origin
- Linear State Estimation
  - Kalman filter
- Nonlinear State Estimation
  - Extended Kalman Filter
  - Deterministic Derivative-free estimators
  - Particle Filters
- Constrained State Estimation
- Estimation under Model-Plant Mismatch
  - Robustness
  - On-line Model Maintenance
- Future research directions

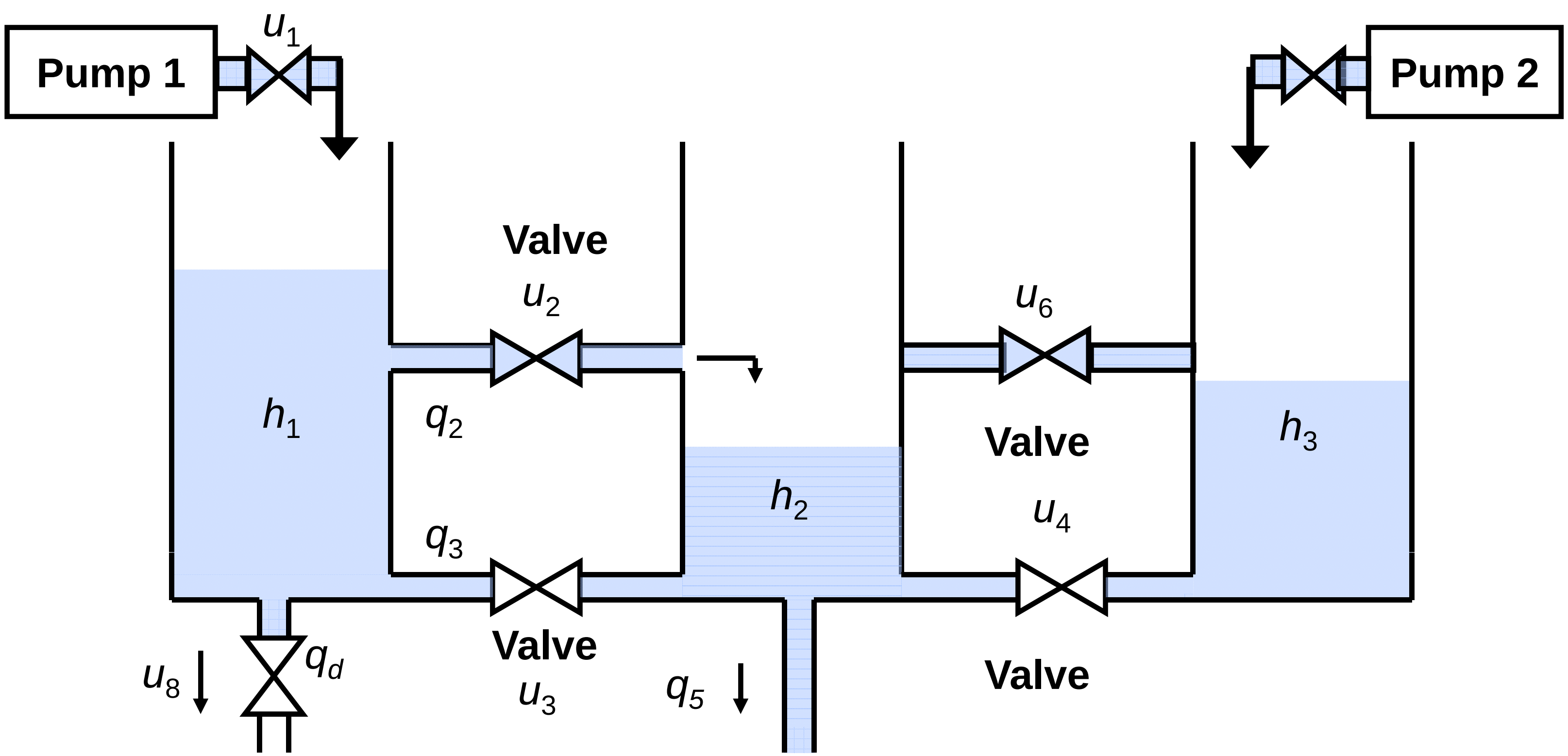
Limitations of EKF

- Covariance update: Using local linearization (Taylor series approximation) of nonlinear system equations
- Requires evaluation of Jacobian at each time step
  - Smoothness requirement on system dynamics: discontinuities not permitted
  - Computationally expensive for large dimension systems
- Propagation step assumes

Mean(Nonlinear Function) ≈ Nonlinear Function (mean)

$$E[g(x)] \approx g[E(x)]$$

Example: Autonomous Hybrid System



Discontinuities in state dynamics: EKF cannot be used

Example: Autonomous Hybrid System

$$A1\frac{dh1}{dt} = q_{max}u_1 - q_2 - q_3 - q_6 \quad (32)$$

$$A2\frac{dh2}{dt} = q_2 + q_3 + q_4 + q_7 - q_5 \quad (33)$$

$$A3\frac{dh3}{dt} = q_{max}u_5 + q_4 + q_7 \quad (34)$$

$$q_2 = z_1 k_2 \sqrt{|(h1' - h2')|}\, u_2$$

$$h1' = h1 - h_T;\ h2' = h2 - h_T$$

$$q_7 = z_2 k_7 \sqrt{|h2' - h3'|}\, u_7$$

$$h2' = h2 - h_T; h3' = h3 - h_T$$

$$z_1 = \begin{Bmatrix} 0 & \{(h1 \le h_T) AND (h2 \le h_T)\} \\ +1 & \left\{\begin{matrix} [(h1 \succ h_T) AND (h2 \le h_T)]\ OR \\ [[(h1 \succ h_T) AND (h2 \succ h_T)] AND (h1' \succ h2')] \end{matrix}\right\} \\ -1 & \left\{\begin{matrix} [(h1 \le h_T) AND (h2 \succ h_T)] OR \\ [[(h1 \succ h_T) AND (h2 \succ h_T)] OR (h1' \prec h2')] \end{matrix}\right\} \end{Bmatrix}$$

State Dependent Discrete Variables

Autonomous Hybrid System

Autonomous Hybrid System: An example of class of systems with discontinuities in the dynamics

The state vector consists of continuous as well as discrete state variables, which can take only integer values.

$$\mathbf{x}(k+1) = F\left[\mathbf{x}(k), \mathbf{u}(k), \xi(k), \mathbf{w}(k)\right]$$
$$\xi(k+1) = G\left[\mathbf{x}(k+1)\right]$$
$$\mathbf{y}(k) = H\left[\mathbf{x}(k)\right] + \mathbf{v}(k)$$

$\xi(k)$: denotes the discrete states and can take only integer values such as -1,0 or +1.

The function G(.) is expressed as a combination of logic variables such as OR, AND, XOR, IF..THEN ..ELSE etc.

Difficulty: Jacobian of F[.] cannot be computed due to discontinuities introduced by the logic variables

Derivative Free Filters

- Basic idea: Better estimates of the moments of a distribution can be obtained using samples rather than using the Taylor series approximation of the nonlinear function (that transforms a random variable)
- Statistical linear regression is used instead of Taylor series approximation

Derivative Free Filters

Deterministic Sampling

- Unscented Kalman Filter
- Divided Difference Filter
- Gauss-Hermite filter
- Central difference filter

Stochastic Sampling

- Particle Filters
- Ensemble Kalman Filter

Statistical Linear Regression

Given a random variable vector, $\mathbf{e}$, and a nonlinear function of the random vector, say $\boldsymbol{\varepsilon} = \mathrm{F}(\mathbf{e})$

by statistical linearization approach a linear approximation of

$$\boldsymbol{\varepsilon} = \mathrm{F}(\mathbf{e}) \cong \mathbf{A}\mathbf{e} + \mathbf{b}$$

is constructed by minimizing $\mathrm{E}\left[\mathrm{e}^{\mathrm{T}}\ \mathrm{e}\right]$

with respect to $(\mathbf{A}, \mathbf{b})$

The optimum is reached for the following choices of (**A**,**b**)

$$\mathbf{A} = \mathbf{P}_{\varepsilon e}\mathbf{P}_{ee}^{-1} \quad and \quad \mathbf{b} = \mathbb{E}\left[\varepsilon\right] - \mathbf{A}\mathbb{E}\left[\mathbf{e}\right]$$

$$\mathbf{P}_{\varepsilon e} = \mathbb{E}\left[\left(\varepsilon - \mathbb{E}\left[\varepsilon\right]\right)\left(\mathbf{e} - \mathbb{E}\left[\mathbf{e}\right]\right)^T\right] \qquad \mathbf{P}_{ee} = \mathbb{E}\left[\left(\mathbf{e} - \mathbb{E}\left[\mathbf{e}\right]\right)\left(\mathbf{e} - \mathbb{E}\left[\mathbf{e}\right]\right)^T\right]$$

Statistical Linear Regression

For these choice of optimal parameters, we have

$$\mathbb{E}\left[\mathbf{e}\right] = \overline{\mathbf{0}}$$

$$\mathbb{E}\left[\mathbf{e}\mathbf{e}^T\right] = \mathbf{P}_{ee}^{-1} - \mathbf{P}_{\varepsilon e}\mathbf{P}_{ee}^{-1}\mathbf{P}_{\varepsilon e}^T$$

Identical expressions for optimal $(\mathbf{A}, \mathbf{b})$
can be derived by minimizing $E\left[\mathbf{e}\mathbf{e}^T\right]$

In the literature, this linear approximation is also referred to as **linear least mean square** (LLMS) estimation.

Statistical Linearization Based Filters

- **Sample generation:** Uses a deterministic sampling technique to select a finite set of sample points

$$\left\{ \hat{\mathbf{x}}^{(i)}(k-1|k-1), \mathbf{w}^{(i)}(k-1), \mathbf{v}^{(i)}(k) : i = 1,2...N \right\}$$

and define associated weights ω_i such that $\sum_{i=1}^{N} \omega_i = 1$

Prediction: Propagate these samples through the system dynamics to compute a cloud of transformed points

Statistical Linearization Based Filters

$$\hat{\mathbf{x}}^{(j)}(k|k-1) = F\left[\hat{\mathbf{x}}^{(j)}(k-1|k-1), \mathbf{u}(k-1), \mathbf{w}^{(j)}(k-1)\right]$$

$$\mathbf{y}^{(j)}(k) = H\left[\hat{\mathbf{x}}^{(j)}(k|k-1)\right] + \mathbf{v}^{(j)}(k)$$

where j = 1,2,...N

Statistics of nonlinearly transformed points

Sample Means

$$\hat{\mathbf{x}}(k|k-1) = \sum_{j=1}^{N} \omega_i \hat{\mathbf{x}}^{(j)}(k|k-1) \qquad \hat{\mathbf{y}}(k|k-1) = \sum_{j=1}^{N} \omega_i \hat{\mathbf{y}}^{(j)}(k|k-1)$$

Statistical Linearization Based Filters

Sample Covariance

$$\mathbf{P}_{\varepsilon,e}(k) \approx \sum_{i=1}^{N} \omega_i \left[\boldsymbol{\varepsilon}^{(i)}(k)\right]\left[\mathbf{e}^{(i)}(k)\right]^T \qquad \mathbf{P}_{e,e}(k) \approx \sum_{i=1}^{N} \omega_i \left[\mathbf{e}^{(i)}(k)\right]\left[\mathbf{e}^{(i)}(k)\right]^T$$

$$\boldsymbol{\varepsilon}^{(i)}(k) = \hat{\mathbf{x}}^{(i)}(k|k-1) - \hat{\mathbf{x}}(k|k-1) \qquad \mathbf{e}^{(i)}(k) = \hat{\mathbf{y}}^{(i)}(k|k-1) - \hat{\mathbf{y}}(k|k-1)$$

Kalman Gain Update

$$L(k) = P_{\varepsilon,e}(k)\left[P_{e,e}(k)\right]^{-1}$$

Updated Mean and Covariance

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \mathbf{L}(k)\left[\mathbf{y}(k) - \hat{\mathbf{y}}(k|k-1)\right]$$

$$P(k|k) = P(k|k-1) - L(k)P_{e,e}(k)L(k)^T$$

Methods for Drawing Samples

Define augmented vector, χ^a, its mean and covariance matrix, $\mathbf{P}^a$, as follows

$$\chi(k-1) = \left[\mathbf{x}(k-1)^T \quad \mathbf{w}(k-1)^T \quad \mathbf{v}(k)^T\right]^T$$

$$\hat{\chi}(k-1|k-1) = \left[\hat{\mathbf{x}}(k-1|k-1)^T \quad \overline{\mathbf{0}}^T \quad \overline{\mathbf{0}}^T\right]$$

$$\mathbf{P}^a(k-1|k-1) = BlockDiag\left[\mathbf{P}(k-1|k-1) \quad \mathbf{Q} \quad \mathbf{R}\right]$$

Sigma Point Generation

$2M+1$ samples are generated where $\mathrm{M} = \dim(\chi)$

$$\chi^{(1)}(k-1|k-1) = \hat{\chi}(k-1|k-1)$$

$$\chi^{(j+1)}(k-1|k-1) = \hat{\chi}(k-1|k-1) + \rho\sqrt{\mathbf{P}_a(k-1|k-1)}\zeta^{(j)}$$

$$\chi^{(j+M+1)}(k-1|k-1) = \hat{\chi}(k-1|k-1) - \rho\sqrt{\mathbf{P}_a(k-1|k-1)}\zeta^{(j)}$$

$$j = 1,2,....,M$$

$\zeta^{(j)}$: Unit vector with j'th element equal to 1 and rest equal to 0

Methods for Drawing Samples

Unscented Kalman Filter (UKF)

$$\rho = \sqrt{M + \kappa} \text{ where } \kappa \text{ is a tuning parameter}$$

$$\omega_1 = \kappa/(M+\kappa) \text{ and } \omega_{i+1} = \omega_{i+M+1} = 1/2(M+\kappa)$$
$$\text{for } i = 1, 2, ..., M$$

These sample points and associated weights have been chosen such that their weighted sample covariance matrix equals $\mathbf{P}^a(k-1|k-1)$

Divided Difference Kalman Filter (DDKF)

Covariance estimate computed using Stirling's multi-dimensional polynomial interpolation.
The first and second order terms in Taylor series approximation are approximated using central difference method with step-size 'h'

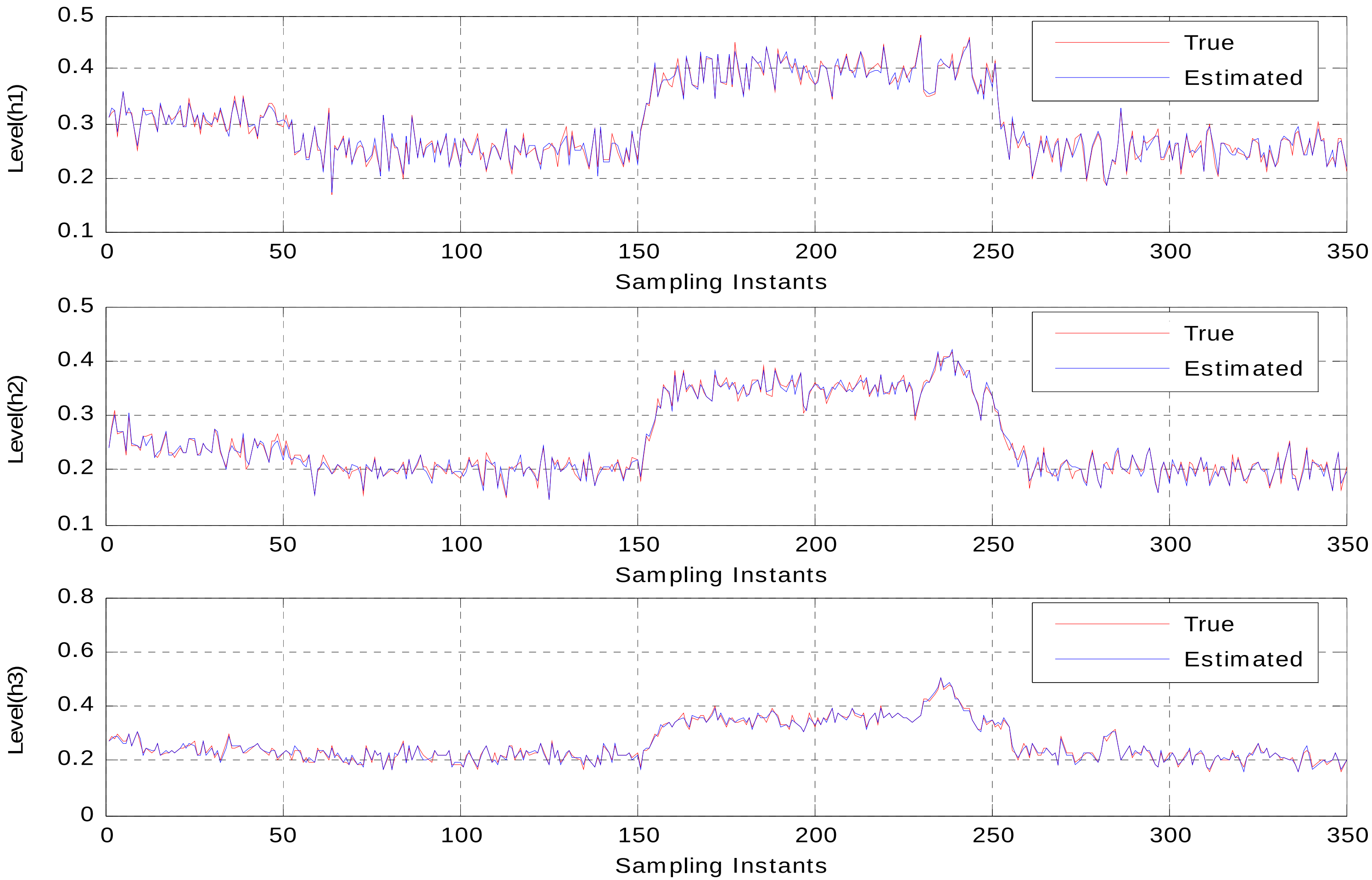
$$\rho = h$$
$$\omega_1 = (h^2 - M)/h^2 \text{ and } \omega_{i+1} = \omega_{i+M+1} = 1/2h^2$$
$$\text{for } i = 1, 2, ..., M$$

Statistical Linearization Based Filters

- UKF results in approximations accurate to third order for Gaussian inputs for all nonlinearities.
- For non-Gaussian inputs, approximations are accurate to at least the second-order. Accuracy of third and higher order moments determined by the choice of tuning parameters
- Sampling based filters be applied for state estimation in systems with discontinuous nonlinear transformations such as autonomous hybrid systems
- **Limitation:** **Do not work well when the conditional densities of states are skewed, Multi-modal, non-Gaussian**

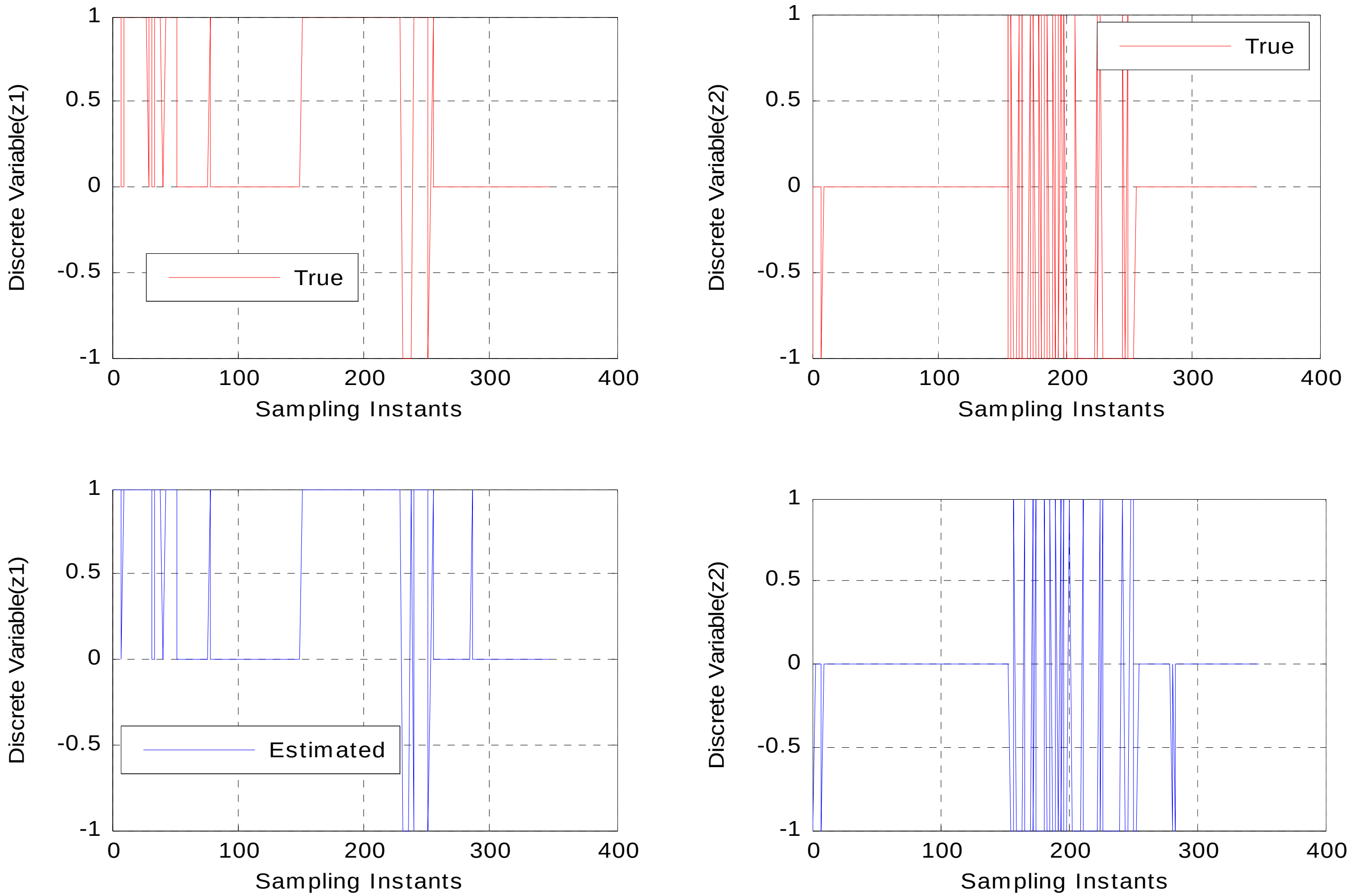
Example: Autonomous Hybrid System

State Estimation using UKF



Example: Autonomous Hybrid System

Discrete State Estimates



Motivating Example

System with skewed, Multi-modal, non-Gaussian conditional densities of states

$$x(k+1) = 0.5x(k) + \frac{25x(k)}{1+x(k)^2} + 8\cos(1.2k) + w(k)$$

$$y(k) = \frac{x(k)^2}{20} + v(k)$$

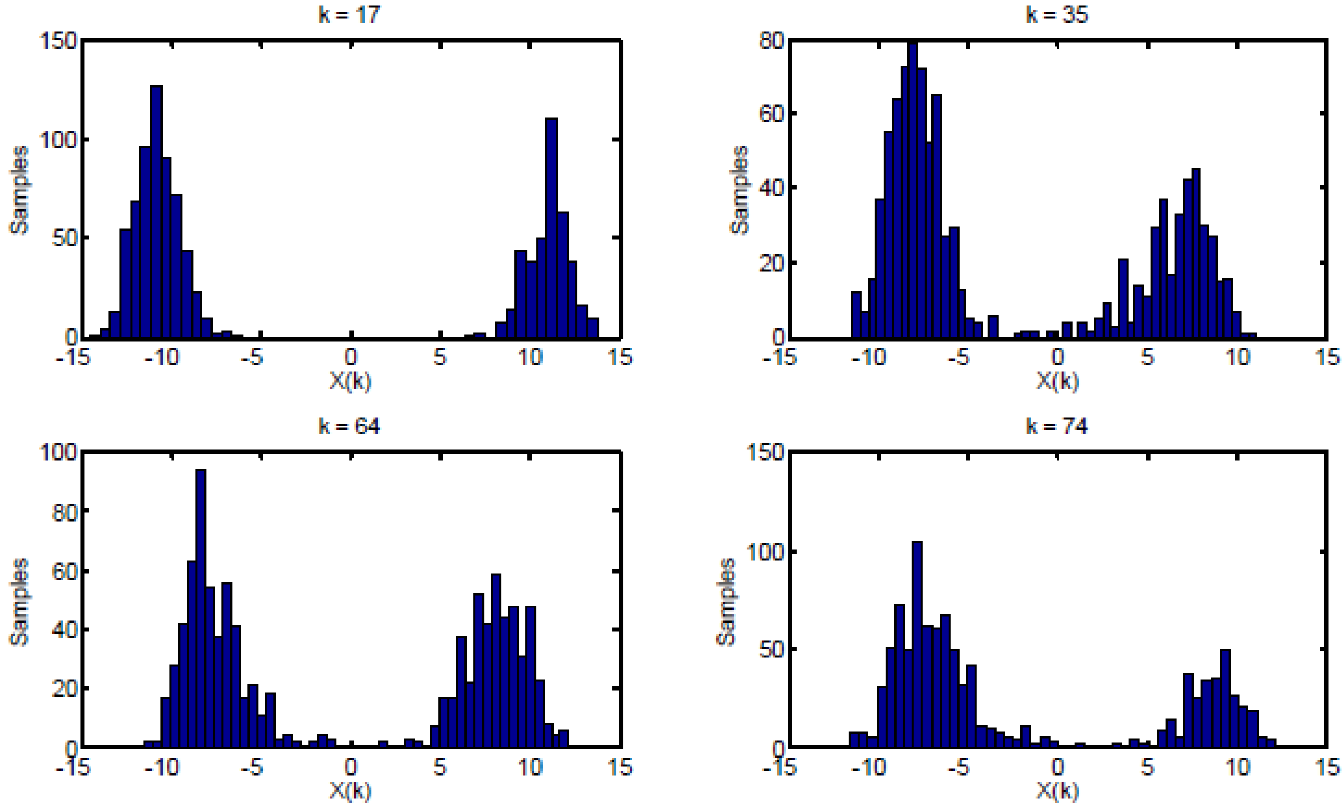
Measurement noise covariance (R) = 1

State noise covariance (Q) = 10

Initial state : x(0|0) = 0 and P(0|0)=10.

Motivating Example

Automation Lab
IIT Bombay



Histogram of particle filter generated samples at few sampling instnat.

Motivating Example

- Histogram indicates that the conditional density of the states are multi-modal and time varying
- Estimators, such as EKF or sigma point filters, which implicitly assume uni-modal conditional densities of states, may not be able to generate accurate state estimates
- This simple example underscores the need to develop better estimation methods for dealing with such pathological systems

Gaussian Sum Filters

- Underlying assumption: any arbitrary PDF can be approximated by a convex combination of Gaussian distributions (Alspach and Sorenson, 1972)

$$p[w(k)] = \sum_{i=1}^{N_w} \beta_i N\left[w, Q^{(i)}\right]$$

- Multiple EKFs are run in parallel
- Updated state estimate: convex combination of individual estimates

$$\hat{\mathbf{x}}(k|k) = \sum_{i=1}^{N_x} \omega_i(k)\hat{\mathbf{x}}^{(i)}(k|k)$$

- Weights recursively updated by application of Bayes' rule and assuming that innovations of individual EKFs have Gaussian distributions.

Gaussian Sum Filters

The conditional densities are approximated via a convex combination of multiple Gaussian densities, i.e.

$$p\left[\mathbf{x}_k|\mathbf{Y}^{k-1}\right] \approx \sum_{j=1}^{L} \mu_{j,k}\mathcal{N}(\widehat{\mathbf{x}}_{k|k}^{(j)}, \mathbf{P}_{k|k-1}^{(j)}) \quad \text{and} \quad p\left[\mathbf{x}_k|\mathbf{Y}^{k}\right] \approx \sum_{j=1}^{L} \mu_{j,k}\mathcal{N}(\widehat{\mathbf{x}}_{k|k}^{(j)}, \mathbf{P}_{k|k}^{(j)})$$

The Gaussian sum assumption implies that

$$\widehat{\mathbf{x}}_{k|k} = \int_{-\infty}^{\infty} \mathbf{x}_k \left[\sum_{j=1}^{L} \mu_{j,k}\mathcal{N}(\widehat{\mathbf{x}}_{k|k}^{(j)}, \mathbf{P}_{k|k-1}^{(j)})\right] d\mathbf{x}_k = \sum_{j=1}^{L} \mu_{j,k}\widehat{\mathbf{x}}_{k|k}^{(j)}$$

Weights are recursively updated by application of Bayes rule

$$\mu_{j,k} = \mu_{j,k-1}\frac{p\left[\mathbf{e}_k^{(j)}|\mathbf{Y}^k\right]}{\sum_{i=1}^{L} p\left[\mathbf{e}_k^{(j)}|\mathbf{Y}^k\right]}$$

where $\left\{\mathbf{e}_k^{(j)}\right\}$ represents the innovation sequence associated with j'th filter.

Particle Filters (PF)

- Can deal with state estimation problems arising from multimodal and non-Gaussian distributions
- Excellent reviews: Arulampalam et al., (2002), Chen, Z. (2003), Bakshi and Rawlings, (2006)
- PF approximates multi-dimensional integration involved in propagation and update steps using Monte Carlo sampling.

Integral:

$$\int_X f(X)p(X)dX = \int_X f(X)dP(X)$$

Approximated as

$$\overline{f}_N = \frac{1}{N}\sum_{i=1}^{N} f\left[X^{(i)}\right]$$

$\{X^{(1)}, X^{(2)},...X^{(N)}\}$: i.i.d. particles drawn from P(X)

Ensemble Kalman Filter

- Proposed by Evensen (1993) and is based on random sampling of the state and the measurement noise from their respective distributions
- Can work with arbitrary distributions of the state disturbance and the measurement noise
- Good combination of stochastic sampling and statistical linearization based filtering: uses only first and second order moments, which are generated using ensemble propagation and update
- Number of samples necessary for generating good estimates can be large

Ensemble Kalman Filter

Computation in prediction and observer gain calculation steps are similar to Statistical Linearization Based Filters. Significant differences are as follows:

1. Samples are drawn only for $\mathbf{w}(k-1)$ and $\mathbf{v}(k)$. Samples of $\mathbf{x}(k-1)$ are propagated from the update step at instant (k - 1)

2. All samples are assigned equal weights, i.e. $\omega_i = 1/N$

3. Update step is used to generate samples for subsequent computations

$$\hat{\mathbf{x}}^{(i)}(k\,|\,k) = \hat{\mathbf{x}}^{(i)}(k\,|\,k-1) + \mathbf{L}(k)\left[\mathbf{y}(k) - \hat{\mathbf{y}}^{(i)}(k|k-1)\right]$$

$$\hat{\mathbf{x}}^{(i)}(k\,|\,k) = \frac{1}{N}\sum_{i=1}^{N}\hat{\mathbf{x}}^{(i)}(k\,|\,k)$$

Thus, no assumption has to be made about the nature of $\mathrm{p}\left[\mathbf{x}(k)\,|\,\mathbf{Y}^{k}\right]$

Particle Filter based on Sequential Importance Sampling

- Difficulty: PDF p(x) of x(k) is unknown

$$E[f(\mathbf{x}(k))] = \int f(\mathbf{x}(k))\, p(\mathbf{x}(k))\, d\mathbf{x}(k)$$

- Solution: Select importance density q(x)

$$E[f(\mathbf{x}(k))] = \int f(\mathbf{x}(k)) \left[\frac{p(\mathbf{x}(k))}{q(\mathbf{x}(k))}\right] q(\mathbf{x}(k))\, d\mathbf{x}(k)$$

Weighting Function

Draw from importance distribution

$$\mathbf{x}^{(i)}(k) \sim q(\mathbf{x})$$

- Draw samples from proposal (importance) distribution
- Weight them according to how they fit the original distribution

Importance Weights

- Particles are weighted by *importance weights.*

$$\overline{f}_N(\mathbf{x}(k)) = \frac{1}{N}\sum_{i=1}^{N}\omega_i f\left[\mathbf{x}^{(i)}(k)\right]$$

Importance weights are updated using Baye's Rule

$$\tilde{\omega}_i(k) = \frac{p\left[\mathbf{x}^{(i)}(1:k) \mid \mathbf{Y}^k\right]}{q\left[\mathbf{x}^{(i)}(1:k) \mid \mathbf{Y}^k\right]}$$

$$= \frac{p\left[\mathbf{y}(k) \mid \mathbf{x}^{(i)}(k)\right] p\left[\mathbf{x}^{(i)}(k) \mid \mathbf{x}^{(i)}(k-1)\right]}{q\left[\mathbf{x}^{(i)}(k) \mid \mathbf{x}^{(i)}(k-1), \mathbf{Y}^k\right]}\tilde{\omega}_i(k-1)$$

$$\omega_i(k) = \frac{\tilde{\omega}_i(k)}{\sum_{j=1}^{N}\tilde{\omega}_j(k)}$$

Importance Sampling using EKF

For example, when EKF is used to generate the importance distribution, steps involved in importance sampling are as follows

For each particle $\mathbf{x}^{(i)}(k-1|k-1)$ where $i = 1,2,...N$

1. Implement EKF and estimate $\bar{\mathbf{x}}^{(i)}(k|k-1), \bar{\mathbf{x}}^{(i)}(k|k)$ and $\mathbf{P}^{(i)}(k|k)$

2. Construct importance density as $\mathbf{N}\left[\bar{\mathbf{x}}^{(i)}(k|k), \mathbf{P}^{(i)}(k|k)\right]$ and draw a new sample from this distribution, i.e. $\mathbf{x}^{(i)}(k|k) \sim \mathbf{N}\left[\bar{\mathbf{x}}^{(i)}(k|k), \mathbf{P}^{(i)}(k|k)\right]$

$$\hat{\mathbf{x}}^{(i)}(k|k) = \bar{\mathbf{x}}^{(i)}(k|k) + [\bar{\mathbf{P}}^{(i)}(k|k)]^{1/2}\Upsilon^{(i)}$$
$$\Upsilon^{(i)} \sim N(0, I)$$

3. Compute the unnormaized weight $\tilde{\omega}_i$ using

$$p\left[\mathbf{y}(k)|\mathbf{x}^{(i)}(k)\right] \equiv \mathbf{N}\left[\mathbf{h}\left(\mathbf{x}^{(i)}(k|k)\right), \mathbf{R}\right]$$
$$p\left[\mathbf{y}(k)|\mathbf{x}^{(i)}(k)\right] \equiv \mathbf{N}\left[\bar{\mathbf{x}}^{(i)}(k|k-1), \mathbf{Q}\right]$$
$$p\left[\mathbf{x}^{(i)}(k)|\mathbf{x}^{(i)}(k-1), \mathbf{Y}^{k}\right] \equiv \mathbf{N}\left[\mathbf{x}^{(i)}(k|k), \mathbf{P}^{(i)}(k|k)\right]$$

Requires running N EKFs in parallel : Computationally Expensive

Particle Filter Algorithm

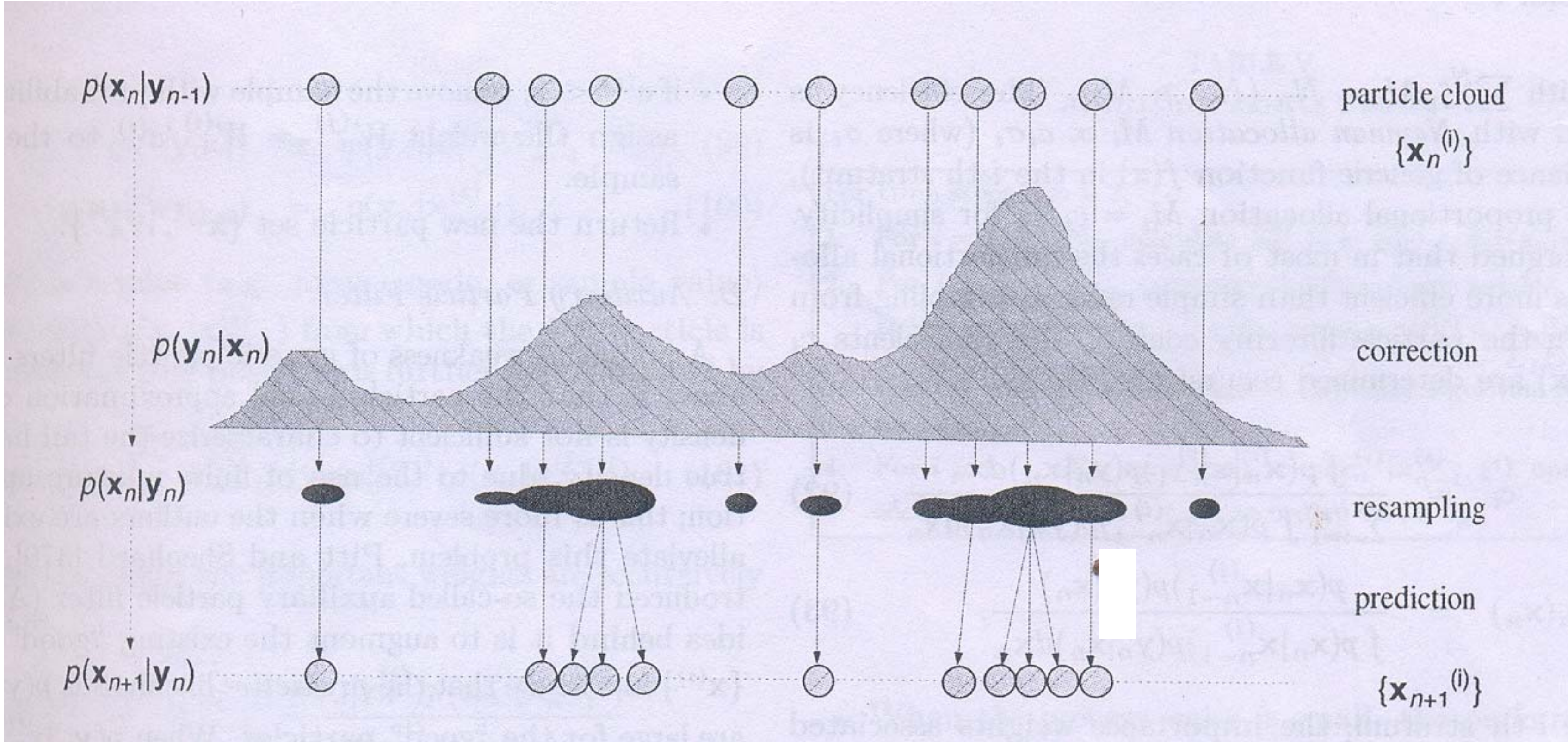
Initialization step: Create particles as samples from the initial state distribution.

As k^{th} instant:

- **Sample** each particle from a proposal distribution
  - EKF as proposal
  - UKF as proposal
- **Compute weights** for each particle using the observation value.
- **Resample particles** generating new particles according to importance weights

PF: Schematic Diagram

PF with importance sampling and re-sampling



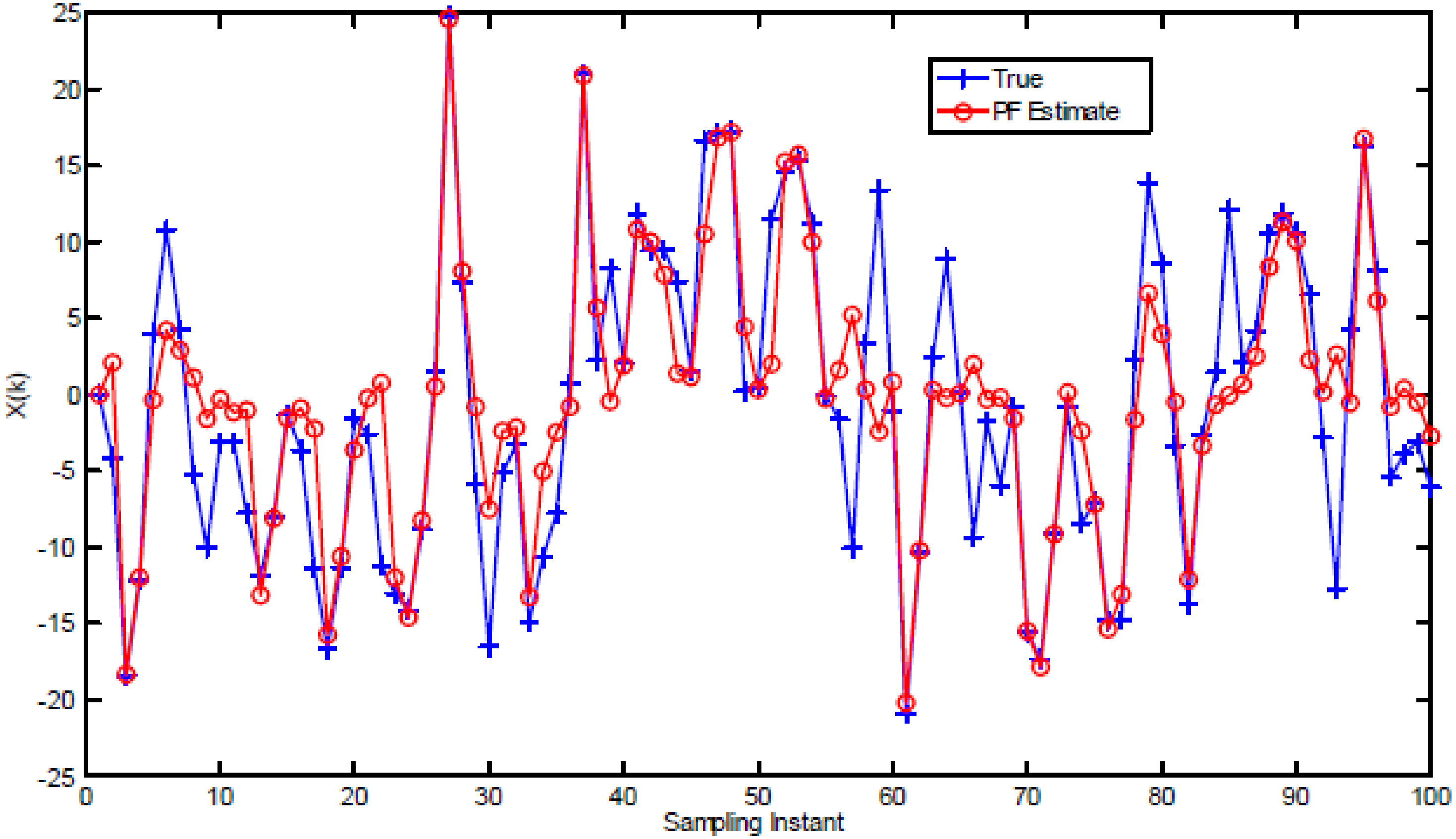
(Figure taken from Chen, Z., 2005)

Motivating Example (Contd.)

Comparison of Mean Sum Squared Estimation Error (SSEE) values over 25 trials

Filter	SSEE	$\frac{SSEE}{SSEE(PF)}$
First Order EKF	79011	37.6
Iterated EKF	19932	9.49
UKF	17727	8.44
EnKF (N = 1000)	8102.5	3.86
SIR PF (N = 1000)	2100	1

Motivating Example (Contd.)



Comparison of true states and states estimated using SIR PF

PF: Advantages & Limitations

- PF can, in principle, deal with arbitrary probability distributions of state propagation error
- It suffers from '**curse of dimensionality**' like most other nonlinear filters developed under Bayesian framework
- Successful when EKF / UKF can be used to generate proposal density. If EKF/UKF diverge, then, most of the samples will be mostly not useful.
- As proposed, it **cannot deal with constraints** on states / parameters

Outline

- Motivation and Origin
- Linear State Estimation
  - Kalman filter
- Nonlinear State Estimation
  - Extended Kalman Filter
  - Deterministic Derivative-free estimators
  - Particle Filters
- Constrained State Estimation
- Estimation under Model-Plant Mismatch
  - Robustness
  - On-line Model Maintenance
- Future research directions

Constrained State estimation

- In most physical systems, **states / parameters are bounded**, which introduces constraints on state / parameter estimates.
  - Moving horizon estimation (**MHE**)
  - Constrained Recursive Formulations
- Moving horizon estimation (**MHE**) (Liebman et al. 1992 , Rao and Rawlings, 2002):
  - State estimation formulated as constrained nonlinear optimization problem over a moving window [k-N:k]
  - Bounds on states/parameters or any other algebraic constraints can be handled

Recursive Constrained Estimation

MHE formulation

- Easy to handle multi-rate and delayed measurements.
- Requires a large dimensional nonlinear optimization problem to be solved at each time step

Recursive constrained formulations

- Based on the premise that the constraint violations occur mostly in the update step
- Combines computational advantages of recursive estimation while handling constraints
- Constrained optimization problem solved over single time step, which make them attractive from the viewpoint of online computations.

Recursive Constrained Estimators

- Constrained EKF or Recursive Dynamic data Reconciliation (RNDDR or C-EKF)

  (Vachhani et al., AIChEJ, 2004)
- Constrained-UKF (C-UKF) or URNDDR

  (Vachhani et al., *Journal of Process Control,* 2006)
- Constrained Ensemble Kalman Filter (C-EnKF)

  (Prakash et. al, I.EC.R., 2010)
- Constrained Particle Filter (C-PF)

  (Prakash et. al, JPC, 2011)

Constrained EKF

RNDDR (or C-EKF)

Prediction Step: State and covariance propagation steps identical to that of EKF

Update Step: solving constrained optimization problem over **[k-1:k]**

$$\hat{\mathbf{x}}(k|k) = \min_{\mathbf{x}(k)} \boldsymbol{\varepsilon}(k|k-1)^T \mathbf{P}(k|k-1)^{-1} \boldsymbol{\varepsilon}(k|k-1) + \mathbf{e}(k)^T \mathbf{R}^{-1} \mathbf{e}(k)$$

$$\boldsymbol{\varepsilon}(k|k-1) = \mathbf{x}(k) - \hat{\mathbf{x}}(k|k-1)$$

$$\mathbf{e}(k) = \mathbf{y}(k) - H[\mathbf{x}(k)]$$

Subject to

$$\mathbf{x}_L \leq \mathbf{x}(k) \leq \mathbf{x}_H$$

Constrained EnKF

Prediction Step: Ensemble prediction identical to that of unconstrained EnKF

Update Step: solving constrained N optimization problems over **[k-1:k]**

$$\hat{\mathbf{x}}^{(i)}(k|k) = \underset{\mathbf{x}(k)}{Min} \left[\boldsymbol{\varepsilon}^{(i)}(k|k-1)^T \mathbf{P}(k|k-1)^{-1} \boldsymbol{\varepsilon}^{(i)}(k|k-1) + \mathbf{e}^{(i)}(k)^T \mathbf{R}^{-1} \mathbf{e}^{(i)}(k) \right]$$

$$\boldsymbol{\varepsilon}^{(i)}(k|k-1) = \mathbf{x}(k) - \mathbf{x}_c^{(i)}(k|k-1) \quad ; \quad \mathbf{e}^{(i)}(k) = \mathbf{y}(k) - \left(H\left[\mathbf{x}(k)\right] + \mathbf{v}^{(i)}(k) \right)$$

Subject to constraints: $\mathbf{x}_L \leq \mathbf{x}(k) \leq \mathbf{x}_U$

$$\overline{\mathbf{x}}(k|k) = (1/N) \sum_{i=1}^{N} \hat{\mathbf{x}}^{(i)}(k|k)$$

Example: Gas-Phase Reaction

Bench-mark Problem: Gas-phase irreversible reaction in a well mixed, constant volume, isothermal batch reactor (Haseltine and Rawlings, 2003)

$$2A \rightarrow B \qquad k_1 = 0.6$$

$$\frac{dp_A}{dt} = -2k_1 p_A^2$$

$$\frac{dp_B}{dt} = k_1 p_A^2$$

$$P = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} p_A \\ p_B \end{bmatrix}$$

$$P(0|0) = \begin{bmatrix} 36 & 0 \\ 0 & 36 \end{bmatrix}$$

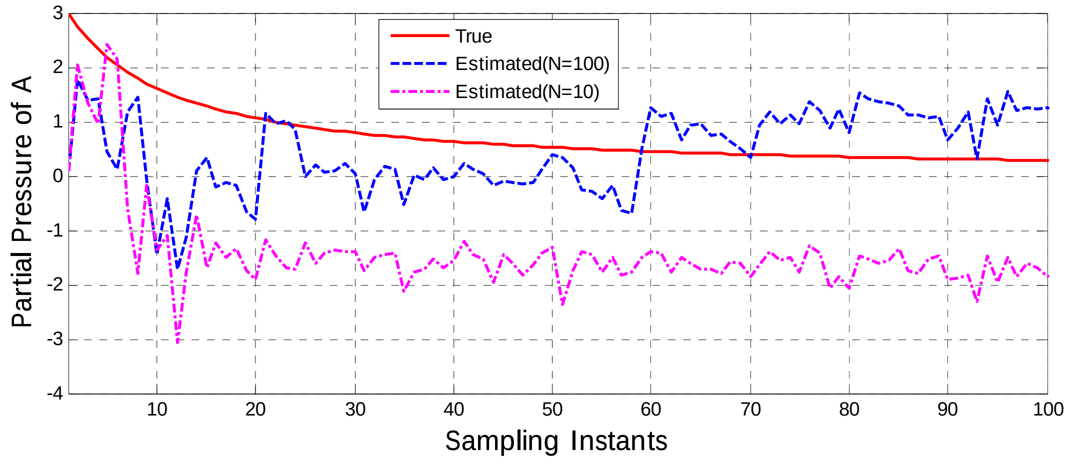
$$\mathbf{x}(0|0) = \begin{bmatrix} 3 & 1 \end{bmatrix}$$

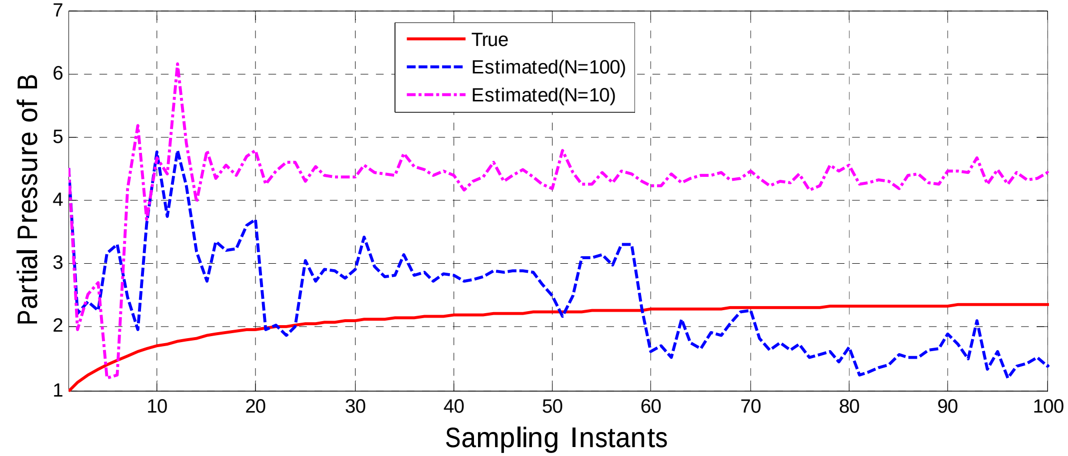
$$\hat{\mathbf{x}}(0|0) = \begin{bmatrix} 0.1 & 4.5 \end{bmatrix}$$

Note that the partial pressures should not be negative.

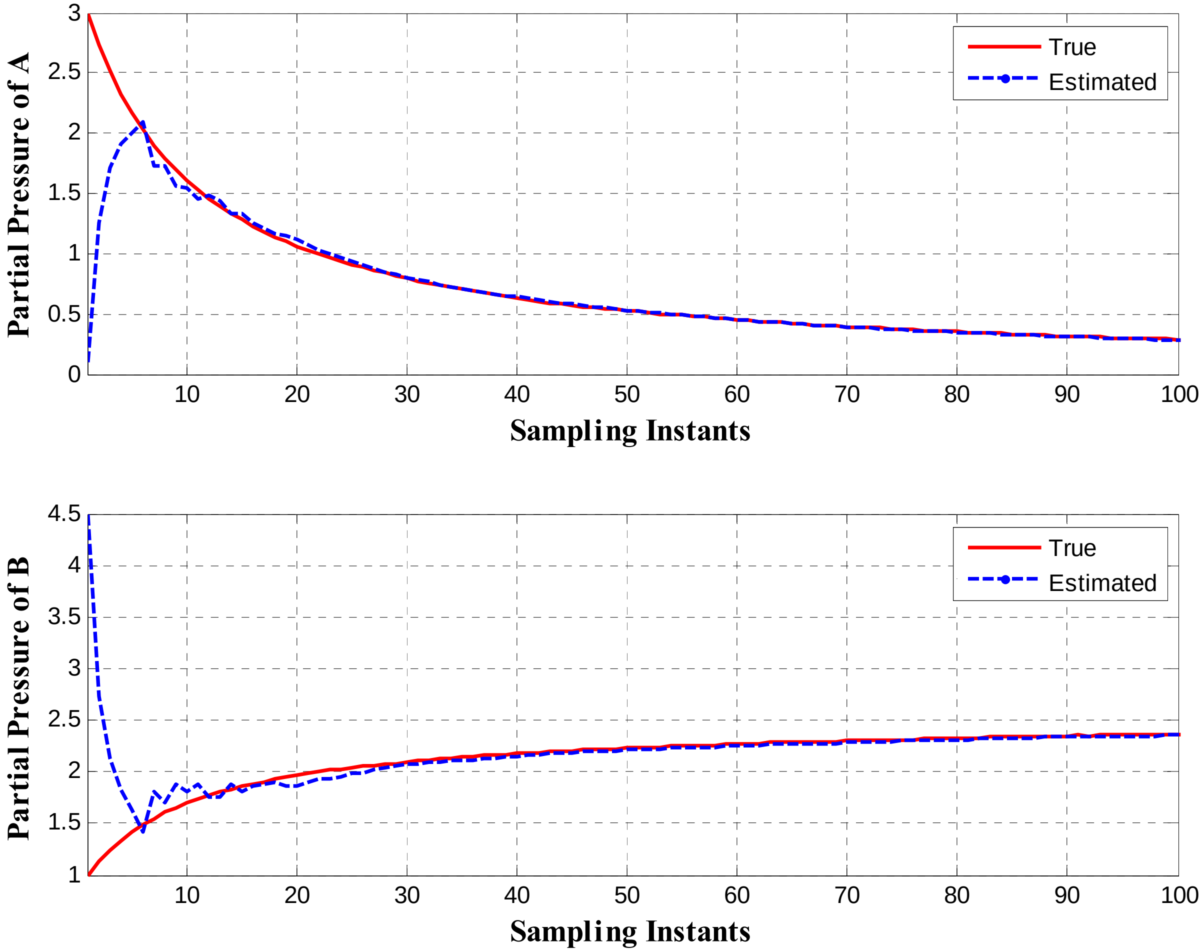
Unconstrained EnKF

Automation Lab
IIT Bombay





Constrained-EnKF



(Prakash, Patwardhan and Shah. *I.E.C.R.*, 2010)

Moving Horizon Estimation

Formulate a sequence of optimization problems over a moving window [k-N,k]

$$\min_{\mathbf{x}(k-N),.....,\mathbf{x}(k)} \left\{ \begin{array}{c} V_{k-N}[\mathbf{x}(k-N)] \\ + \sum_{j=k-N}^{k-1} [\mathbf{w}(j)]^T \mathbf{Q}^{-1}\mathbf{w}(j) + \sum_{j=k-N}^{k} [\mathbf{v}(j)]^T \mathbf{R}^{-1}\mathbf{v}(j) \end{array} \right\}$$

Subject to

$$\mathbf{w}(\mathrm{j}) = \mathbf{x}(\mathrm{j}+1) - \mathbf{F}[\mathbf{x}(\mathrm{j}), \mathbf{u}(j)]$$

$$\mathbf{v}(\mathrm{j}) = \mathbf{y}(\mathrm{j}) - \mathbf{H}[\mathbf{x}(\mathrm{j})]$$

$$\text{Bounds on state}: \mathbf{x}_L \leq \mathbf{x}(j) \leq \mathbf{x}_H$$

Solution yields smoothed and current state estimates,

i.e., $\hat{\mathbf{x}}(\mathrm{k-N} \mid \mathrm{k}), \hat{\mathbf{x}}(1 \mid \mathrm{k}),....., \hat{\mathbf{x}}(\mathrm{k-1} \mid \mathrm{k}), \hat{\mathbf{x}}(\mathrm{k} \mid \mathrm{k})$

under the constraints (bounds on the states)

Moving Horizon Estimation

$V_{k-N}[\mathbf{x}(k-N)]$: Arrival Cost

$$V_{k-N}[\mathbf{x}(k-N)] = \left\{ \begin{array}{c} \|\mathbf{x}(0)-\hat{\mathbf{x}}(0|0)\|^2_{\mathbf{P}(0|0)^{-1}} \\ + \sum_{j=1}^{k-N-1} [\mathbf{w}(j)]^T \mathbf{Q}^{-1} \mathbf{w}(j) + \sum_{j=0}^{k-N-1} [\mathbf{v}(j)]^T \mathbf{R}^{-1} \mathbf{v}(j) \end{array} \right\}$$

$$= -\log p\left[\mathbf{x}(k-N) \mid \mathbf{Y}^{k-N}\right]$$

Important to construct reasonably accurate estimates of the Arrival Cost: an open issue in MHE literature

Conditional density : difficult to estimate in the constrained nonlinear case.

Arrival Cost Estimation

Automation Lab
IIT Bombay

Suppose we decide to use the mean squared error type approximation for the arrival cost, i.e.

$$V_{k-N}\left[\mathbf{x}(k-N)\right] \approx \left\| \mathbf{x}(k-N) - \hat{\mathbf{x}}(k-N \mid k-N) \right\|^2_{\mathbf{P}(k-N|k-N)^{-1}}$$

then, how to estimate $\mathbf{P}(k-N \mid k-N)$ in the presence of consraints?

- Constrained (sampling based) Recursive Bayesian estimators, such as C-EnKF or C-PF, are better suited for arrival cost estimation
- Covariance estimate generated from the constrained samples is used for approximating the arrival cost.

(López-Negrete et al., JPC, 2011)

Arrival Cost Estimation: Case Study

CSTR System with constraints on states

$$\frac{dC_A}{dt} = \frac{F}{V}(C_A^{in} - C_A) - 2k(T_R)C_A^2$$

$$\frac{dT_R}{dt} = \frac{F}{V}(T_R^{in} - T_R) + \frac{2(-\Delta H_R)k(T_R)C_A^2}{\rho C_P} - \frac{UA}{V\rho C_p}(T_R - T_{cw})$$

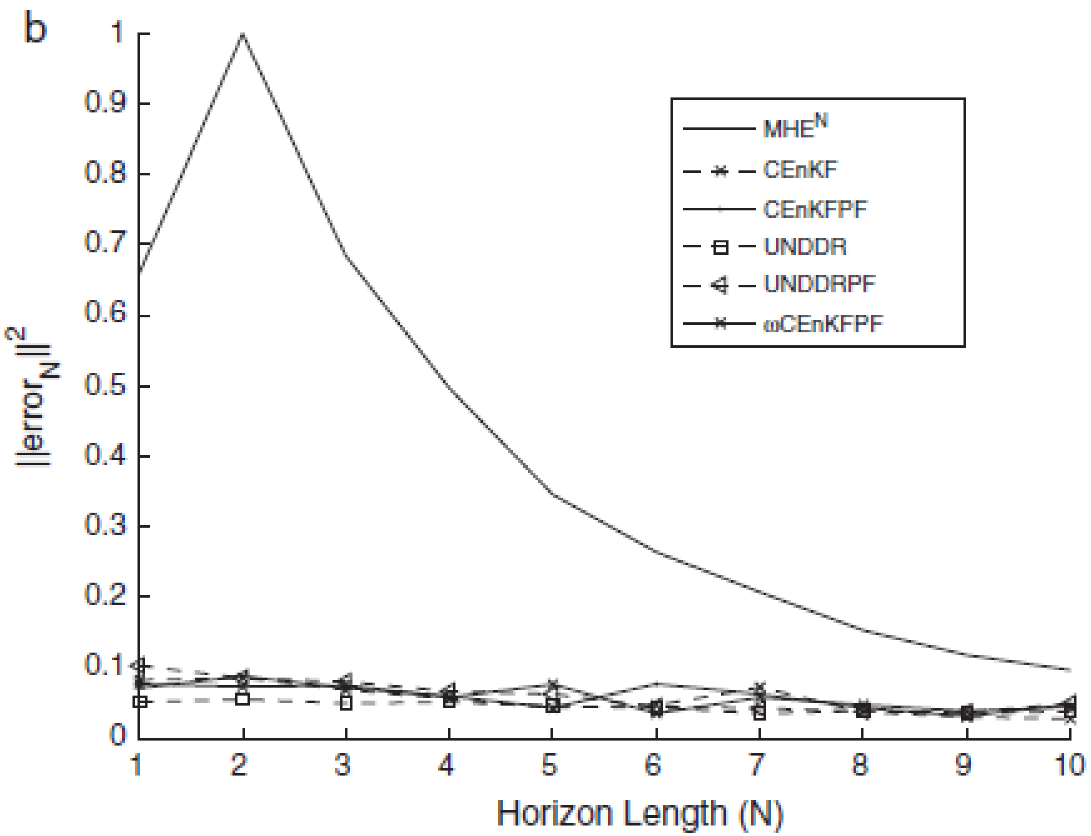
$$\frac{dT_{cw}}{dt} = \frac{F_{cw}}{V_{cw}}(T_{cw}^{in} - T_{cw}) + \frac{UA}{V_{cw}\rho_{cw}C_{pcw}}(T_R - T_{cw})$$

$$k(T_R) = k_o \exp\left[\frac{-E_a}{RT_R}\right]$$

$$C_A \in [0, 1],\ T_R \in [200, 420],\ T_{cw} \in [200, 420].$$

If arrival cost is estimated with a constrained recursive filter instead of EKF, can it reduce the window size and, in turn, reduce the on-line computations?

Arrival Cost Estimation: Case Study



MSE as a function of horizon length when using constrained filters for the arrival cost approximation for the CSTR example

Arrival Cost Estimation: Case Study

- EKF based approximations of the arrival cost introduce unwanted errors, which require the choice of longer horizon lengths and a larger optimization problem to be solved on-line.
- Particle-based filters can approximate arrival cost distributions using samples, and thus require few assumptions on the type of distribution. Moreover, CEnKF and constrained C-PF handle bounds on the states, and thus provide a more consistent approximation of the arrival cost.
- Resulting improvements in the arrival cost approximation allow us to use a smaller horizon window for MHE, and a smaller NLP can be solved on-line

Outline

- Motivation
- Linear State Estimation
  - Kalman filter
- Nonlinear State Estimation
  - Extended Kalman Filter
  - Deterministic Derivative-free estimators
  - Particle Filters
- Constrained State Estimation
- **Robust Estimation and On-Line Model Maintenance**
- Future research directions

Estimation with M-P-M

- Model-Plant-Mismatch
  - Parameter drifts / abrupt changes
    - Equipment fouling
    - Catalyst degradation
    - Leaks
  - Sensor / actuator biases / failures
- Is the state estimator “robust” to MPM?
  - Is estimation error bounded if MPM is bounded?
- Can we find which part of the model is bad ?

Robustness of EKF

- Extended the nominal convergence proof of by Reif et al. (1999) to show

  "If MPM is restricted to a compact set, then the observer errors are bounded (i.e. input to state (ISS) stable)"

- Using a EKF in Nonlinear MPC for offset free control:

  "If observer is ISS and NMPC is nominally stable, then closed loop system obtained by combining the observer with the NMPC is Input-to-State practically Stable (ISpS)"

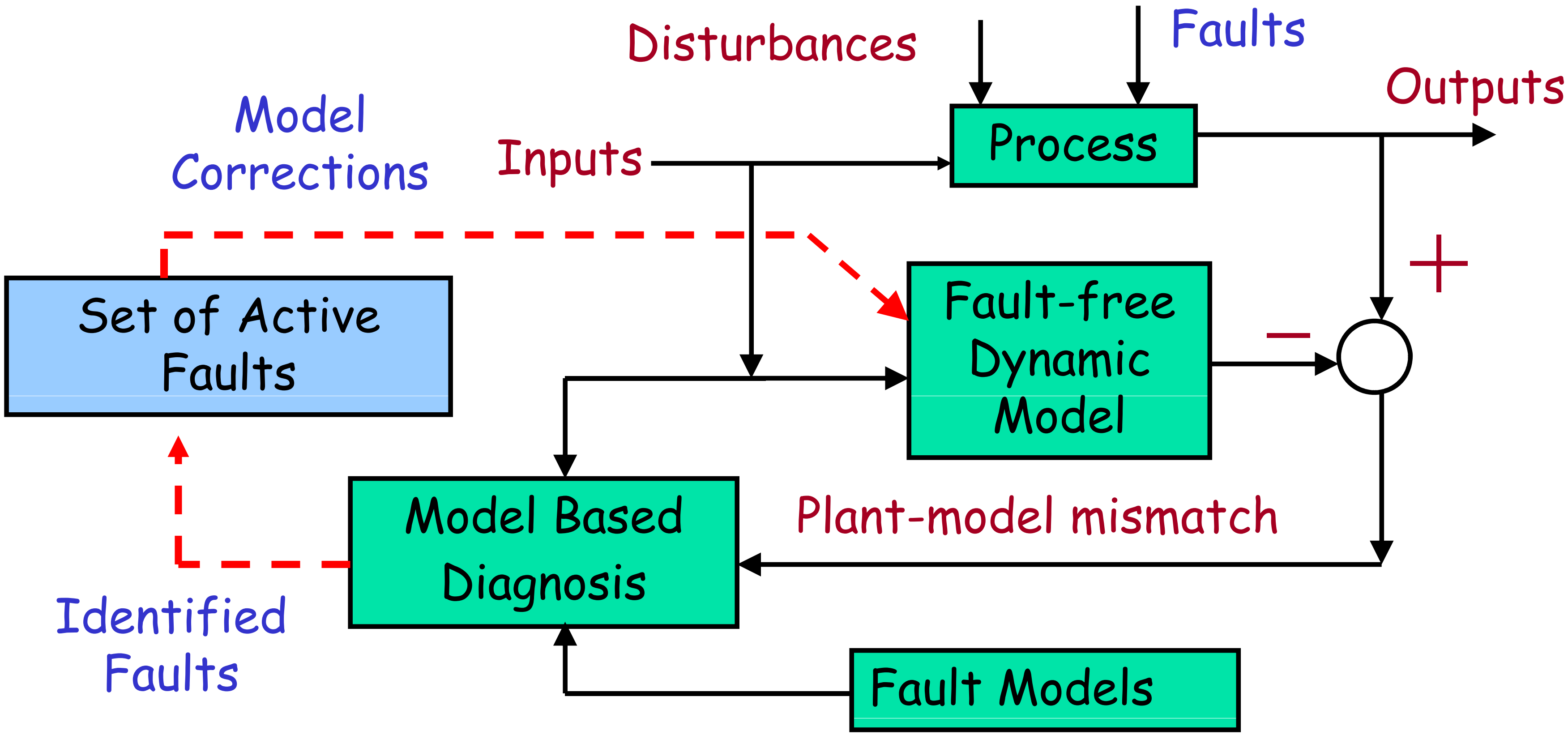
(Huang, Patwardhan, Biegler, JPC, 2011)

Simultaneous State & Parameter Estimation

- Direct approach (state augmentation)
  - Augment state vector with extra states corresponding to faults
  - Simultaneously estimate state and 'fault states'
- Advantages
  - Arbitrary type of fault behavior (step/slow drift) can be tracked
  - Magnitude estimate of the fault is available and can be used for achieving fault tolerance
- Limitation
  - Number of extra states which can be estimated cannot exceed number of measurements.

Active Model Maintenance

Fault Diagnosis: Sophisticated schemes for one-time abnormal behavior identification



(Deshpande at al., JPC, 2009)

Active Model Maintenance

- Objectives
  - Online detection of multiple *abrupt changes* occurring sequentially in time
  - On-line model correction based on diagnosis
- Approach: GLR Method
  - Diagnosis: Generalized Likelihood Ratio method
  - Innovation sequences generated by KF / EKF carry signature of change
  - Exploits the pattern of innovation to identify fault magnitude fault type
  - Fault that corresponds to maximum value of ***likelihood ratio*** is identified as fault

Issues in State Estimation

- Robustness to plant-model mismatch: Model accuracy is critical to state estimation
- Noise Model Parameters: Measurement and state noise co-variances are difficult to estimate. These matrices are often treated as tuning parameters
- Number of extra states (unmeasured disturbances / parameters) estimated cannot exceed number of measurements
- Computationally efficient methods for irregularly sampled multi-rate measurement scenario
- Conditional density and arrival cost estimation in presence of constraints on states

Research Directions

Nonlinear state estimation: rich and highly active research area

- State estimation of systems governed by DAE in Bayesian framework
- Sampling based filters suffer from the **curse of dimensionality**: handling computational complexities that can arise in large scale systems
- Optimal state estimation in the presence of inequality constraints
- Integration of fault diagnosis techniques for on-line model maintenance: isolation of active set of changing parameters and dealing with structural MPM
- Estimation in presence of irregularly sampled and delayed measurements
- Simultaneous state and parameter estimation

Acknowledgements

- Collaborators
  - Prof. Sirish Shah (Univ. of Alberta, Canada)
  - Prof. S. Narasimhan (IIT, Madras)
  - Prof. Lorenz Biegler (CMU)

- Graduate Students
  - Dr. J. Prakash (IIT Madras)
  - Dr. Seema Manuja (IIT Madras)
  - Dr. Anjali Deshpande (IIT Bombay)
  - Dr. Abhijit Badwe (IIT Bombay)
  - Dr. Vinay Bavdekar (IIT Bombay)
  - Dr. Rui Huang (CMU)
  - Dr. Rodrigo López-Negrete (CMU)

Acknowledgements

- Board of Research for Nuclear Sciences, Govt. of India
- NSERC-Matrikon-Suncore-Icore Industrial Research Chair program in Computer Process Control at University of Alberta
- Indo-U.S. Science and Technology Forum (IUSTF) travel grant

Thank You !

Questions ?

References

Review Articles / Books

- Arulampalam, S. Maskell, N. Gordon and T. Clapp, (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, *IEEE Transactions on Signal Processing,* **50(2)**, pp. 174–188.
- Burgers, G, Leeuwen, P. J. V., Evensen G. (1998). Analysis Scheme in the Ensemble Kalman Filter, Monthly Weather Review, 126, 1719-1724.
- Chen, Z. (2003). Bayesian filtering: from Kalman filters to particle filters, and beyond, Technical Report, Adaptive Syst. Lab., McMaster University, Hamilton, ON, Canada.
- Daum, F. (2005). Nonlinear filters: Beyond the Kalman filter, *IEEE A&E Systems Magazin,* **20**, **(8)**, 57–69.
- Julier S.J. and J.K. Uhlmann (2004), Unscented filtering and nonlinear estimation, *Proceedings of the IEEE,* **92** (3), pp. 401–422.
- Muske, K.R. and T.F. Edgar **(1997).** Nonlinear state estimation. In *Nonlinear Process Control* (M.A. Henson and D.E. Seborg, Eds.), Prentice-Hall.
- Soderstorm,T.(2002), Discrete-time stochastic systems, Advanced Textbooks in Control and Signal Processing, Springer.
- Sorenson, H. (1985). Kalman Filtering: Theory and Applications, IEEE Press, New York, 1985.

References

Research Papers in referred in the slides

- Alspach, D. L., Sorenson, H. W. (1972). Nonlinear Bayesian Estimation Using Gaussian Sum Approximations, IEEE Transactions on Automatic Control, 4, 439-448.
- Bavdekar, V. A., Deshpande, A. P., Patwardhan, S. C. (2011a). "Identification of process and measurement noise covariances for state and parameter estimation using extended Kalman filter", Journal of Process Control, 21, 585-601.
- Deshpande, A., Patwardhan, S. C., Narasimhan, S. Intelligent State Estimation for Fault Tolerant Nonlinear Model Predictive Control, *Journal of Process Control,* 19, 187–204, **2009**.
- Huang, R., Patwardhan, S. C., Biegler, L. T. (2010). Stability of a class of discrete-time nonlinear recursive observers, J. of Process Control, 20, 1150--1160.
- López-Negrete, R.; Patwardhan, S.C.; and Biegler, L.T., (2011). Constrained particle filter approach to approximate the arrival cost in Moving Horizon Estimation. Journal of Process Control, 21, pp. 909-919.
- Prakash, J., Patwardhan, S. C., Shah, S. L On the choice of importance distributions for unconstrained and constrained state estimation using particle filter, Journal of Process Control, 21, pp. 119-129.

References

- Patwardhan, S. C., Narasimhan, S., Prakash, J., Gopaluni, B., Shah, S. L (2012) Nonlinear Bayesian State Estimation: A Review Of Recent Developments". *Accepted for publication in Control Engineering Practice, 2012.*
- Prakash, J., S. C. Patwardhan and S. L. Shah (2010) "Constrained State Estimation Using the Ensemble Kalman Filter",, *Ind. Eng. Chem. Res., 49 (5), pp 2242-2253.*
- Rawlings, J. B., and B R. Bakshi, Particle filtering and moving horizon estimation, *Computers and Chemical Engineering*, 30, 10-12, 1529-1541, 2006.
- Reif, K., Gunther, S., Yaz, E., Unbehauen, R. (1999). Stochastoc stability of the discrete-time extended Kalman filter. IEEE Trans. on Automatic Control, 44,4, 714-728.
- Sorenson, H. (1970). Least square estimation from Gauss to Kalman. IEEE Spectrum, pp 63-68, July, 1970.
- Vachhani, P., R. Rengaswamy, V. Gangwal and S. Narasimhan (2004), Recursive estimation in constrained nonlinear dynamical systems, *AIChE J.*, 946–959.
- Vachhani, P., Narasimhan, S., Rengaswamy, R. (2006). Robust and reliable estimation via Unscented Recursive Nonlinear Dynamic Data Reconciliation. Journal of Process Control, 16, 1075-1086.