Mini-Course 07
Kalman Particle Filters

Henrique Massard da Fonseca

Cesar Cunha Pacheco

Wellington Bettencurte

Julio Dutra

Agenda

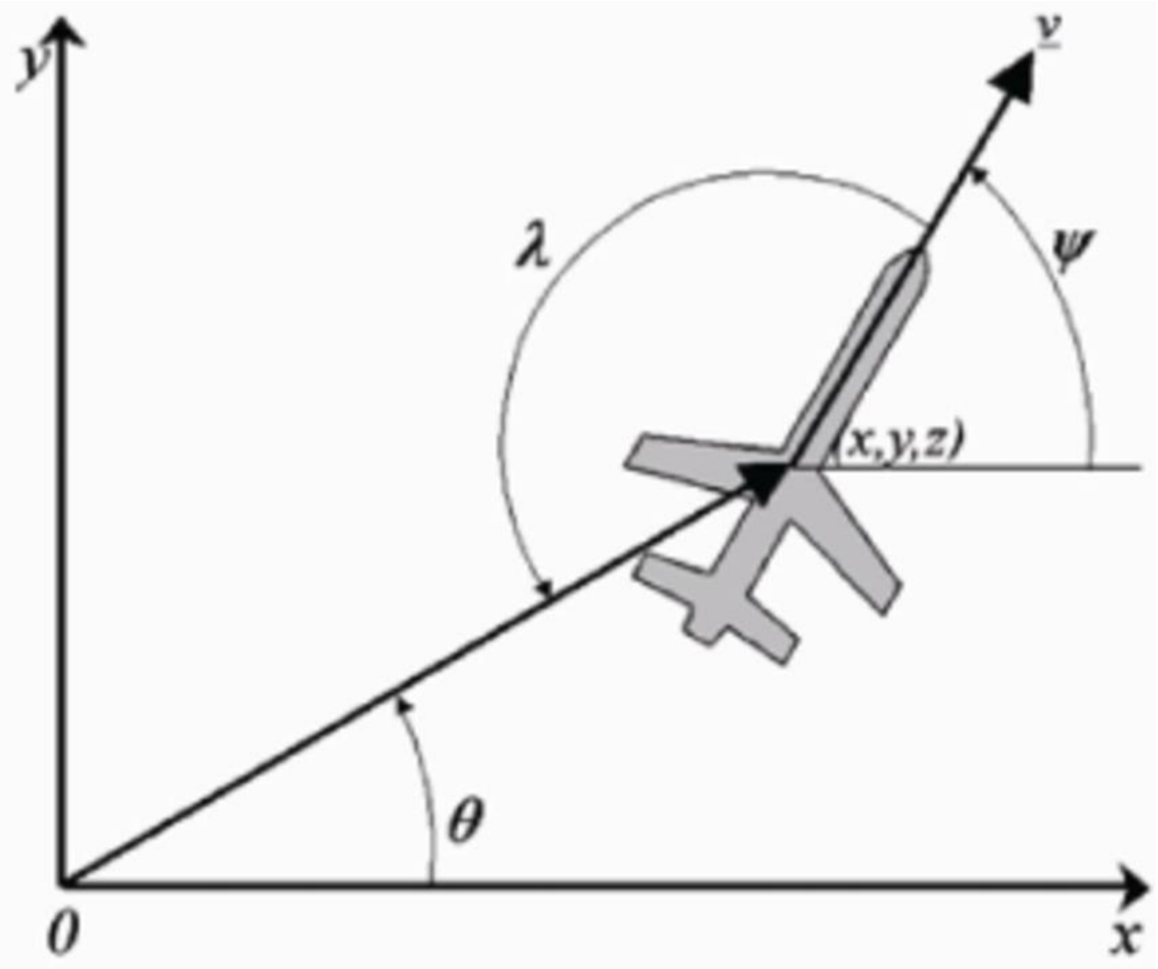
- State Estimation Problems & Kalman Filter
  - Henrique Massard
- Steady State Kalman Filter, Extended & Encented Kalman Filter
  - Cesar Cunha Pacheco
- Particle Filter
  - Wellington Bettencurte and Julio Dutra

State Estimation Problem

- Introduction
  - State **estimation problems**, also designated as **nonstationary inverse problems;**
  - Available **measured data** is used together with **prior knowledge** about the **physical phenomena** and the **measuring devices**, in order to **sequentially** produce **estimates** of the desired **dynamic variables**;
  - This is accomplished in such a manner that the error is minimized **statistically**.

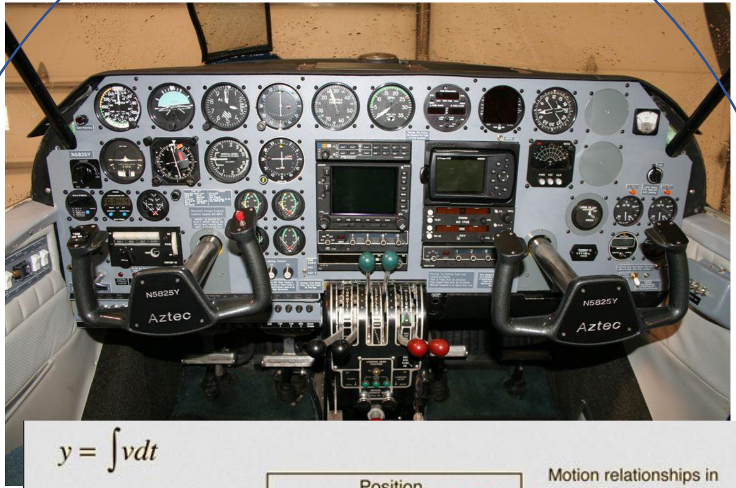
State Estimation Problem

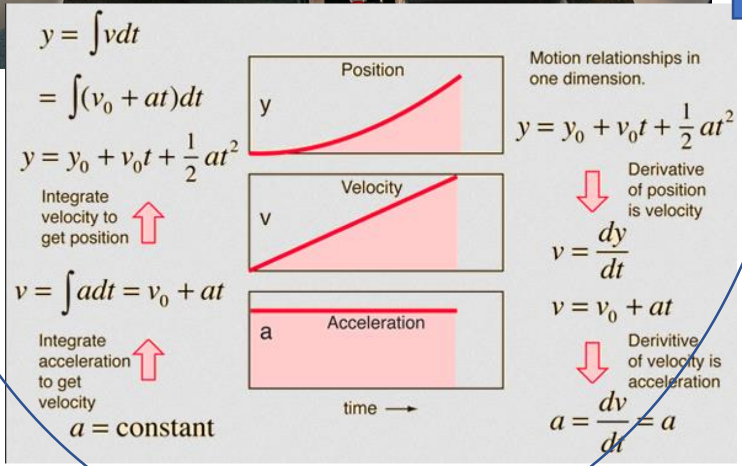
- Position of an aircraft
  - Time integration of velocity components since departure
    - Models aren´t perfect
  - Measured with an GPS and altimeter
    - Measurement devices aren´t perfect



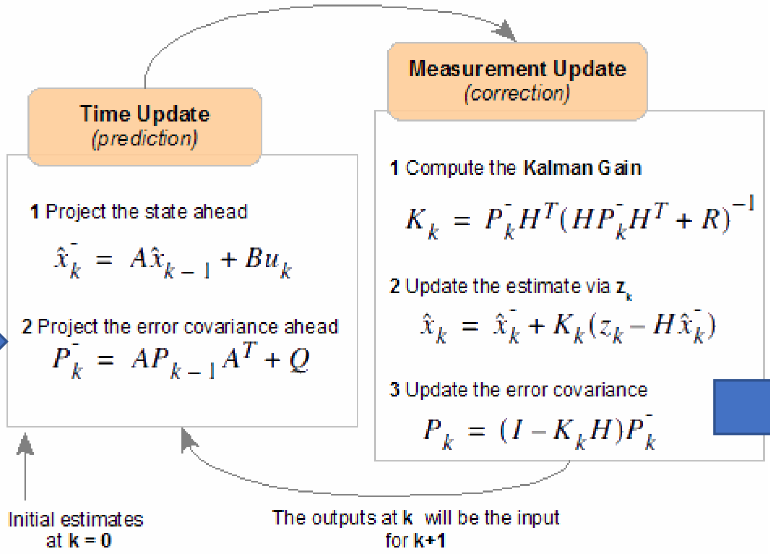
State Estimation Problem

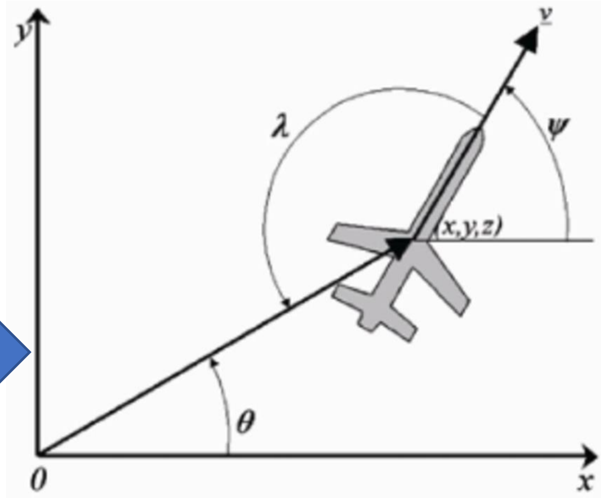
Measurement error





Model error





State Estimation Problem

- Consider the model for the evolution of the vector $\mathbf{x}$

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{w}_{k-1})$$

- Measurements available and are related to $\mathbf{x}_k$ as

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{v}_k)$$

k=1,2,... denotes time instant

- Where:

$\mathbf{x} \in R^{n_x}$ = state variables to be estimated

$\mathbf{y} \in R^{n_y}$ = measurements

$\mathbf{w} \in R^{n_w}$ = state noise vector

$\mathbf{v} \in R^{n_v}$ = measurement noise

State Estimation Problem

- The state estimation problem aims at obtaining information about $\mathbf{x}_k$ based on the **state evolution model** and on the **measurements** given by the **observation model**.

Bayesian Framework

- The **solution** of the **inverse problem** within the **Bayesian framework** is recast in the form of statistical inference from the **posterior probability density**, which is the **model for the conditional probability distribution of the unknown parameters given the measurements.**
  - The **measurement model** incorporating the related uncertainties is called the **likelihood**, that is, the **conditional probability of the measurements given the unknown parameters**.
  - The **model for the unknowns** that reflects all the uncertainty of the parameters **without** the **information** conveyed by the measurements, is called the **prior** model.

Bayesian Framework

- The formal mechanism to combine the new information (measurements) with the previously available information (prior) is known as the <u>**Bayes' theorem:**</u>

$$\pi_{posterior}(\mathbf{x}) = \pi(\mathbf{x}|\mathbf{y}) = \frac{\pi(\mathbf{x})\pi(\mathbf{y}|\mathbf{x})}{\pi(\mathbf{y})}$$

where $\pi_{posterior}(\mathbf{x})$ is the **posterior probability density**, $\pi(\mathbf{x})$ is the **prior density**, $\pi(\mathbf{y}|\mathbf{x})$ is the **likelihood function** and $\pi(\mathbf{y})$ is the **marginal probability density** of the measurements, which plays the role of a normalizing constant.

State Estimation Problem

- **State Evolution Model:** $\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{w}_{k-1})$
- **Observation Model:** $\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{v}_k)$
- The *evolution-observation model* is based on the following assumptions:
  - The sequence $\mathbf{x}_k$ for k=1,2…., is a Markovian process, that is,

$$\pi(\mathbf{x}_k|\mathbf{x}_0,\mathbf{x}_1,\mathbf{x}_2,\ldots\mathbf{x}_{k-1}) = \pi(\mathbf{x}_k|\mathbf{x}_{k-1})$$

  - The sequence $\mathbf{y}_k$ for k=1,2,…, is a Markovian process with respect to the history of $\mathbf{x}_k$, that is,

$$\pi(\mathbf{y}_k|\mathbf{x}_0,\mathbf{x}_1,\mathbf{x}_2,\ldots\mathbf{x}_k) = \pi(\mathbf{y}_k|\mathbf{x}_k)$$

  - The sequence $\mathbf{x}_k$ depends on the past observations only through its own history, that is,

$$\pi(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{y}_{1:k-1}) = \pi(\mathbf{x}_k|\mathbf{x}_{k-1})$$

State Estimation Problem

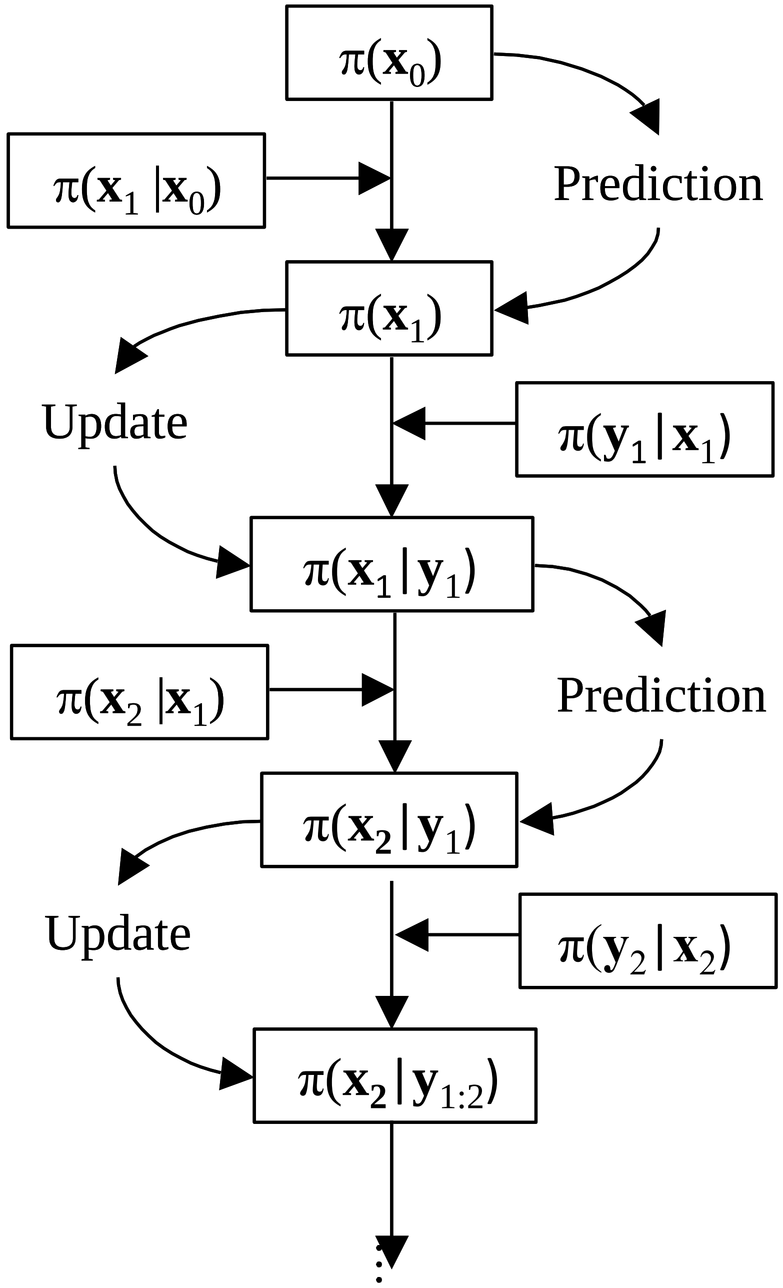
State Evolution Model: $\mathbf{x}_k = \mathbf{f}_k\,(\mathbf{x}_{k-1},\mathbf{w}_{k-1})$

Observation Model: $\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{v}_k)$

- *time* domain: used to classify the method of solution in the domain $\mathbf{y}_{1:kf}=\{\mathbf{y}_k, k=1,\ldots,kf\}$
  - The **prediction problem**, concerned with the determination of $\pi(\mathbf{x}_k|\mathbf{y}_{1:k-1})$;
  - The **filtering problem**, concerned with the determination of $\pi(\mathbf{x}_k|\mathbf{y}_{1:k})$;
  - The **fixed-lag smoothing problem**, concerned with the determination of $\pi(\mathbf{x}_k|\mathbf{y}_{1:k+p})$, where $p\geq 1$ is the fixed lag;
  - The **whole-domain smoothing** problem, concerned with the determination of $\pi(\mathbf{x}_k|\mathbf{y}_{1:kf})$

State Estimation Problem

- Filtering Problem
  - By assuming that $\pi(\mathbf{x}_0 | \mathbf{y}_0) = \pi(\mathbf{x}_0)$ is available, the posterior probability density $\pi(\mathbf{x}_k | \mathbf{y}_{1:k})$ is then obtained with Bayesian filters in two steps: **prediction and update**



The Kalman Filter

- Evolution and observation models are linear.
- Noises in such models are additive and Gaussian, with known means and covariances.

State Evolution Model: $\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{s}_{k-1} + \mathbf{w}_{k-1}$

Observation Model: $\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k$

- The above set o equations leads to the optimum solution of the state variables;
- $\mathbf{F}$ and $\mathbf{H}$ are known matrices for the linear evolutions of the state $\mathbf{x}$ and of the observation $\mathbf{y}$, respectively.
- Vector $\mathbf{s}$ is assumed to be a known input .
- Noises $\mathbf{w}$ and $\mathbf{v}$ have zero means and covariance matrices $\mathbf{Q}$ and $\mathbf{R}$, respectively – $p(w) \sim N(0,Q)$ and $p(v) \sim N(0,R)$

The Kalman Filter

- Derivation of the equations
  - The inverse problem solution can be obtained by

$$\widehat{\mathbf{x}}_n = \max_{\mathbf{x}_n} \{\pi(\mathbf{x}_n|\mathbf{y}_{0:n})\}$$

  - Where the posterior pdf can be obtained by the following proportionality relation:

$$\pi(\mathbf{x}_n|\mathbf{y}_{0:n}) \propto \pi(\mathbf{y}_n|\mathbf{x}_n)\pi(\mathbf{x}_n|\mathbf{y}_{0:n-1})$$

  - Reminding the evolution observation models:

$$\mathbf{x}_n = \mathbf{F}_n\mathbf{x}_{n-1} + \mathbf{w}_{n-1}$$

$$\mathbf{y}_n = \mathbf{H}_n\mathbf{x}_n + \mathbf{v}_n$$

The Kalman Filter

- Derivation of the equations

  - Determine the **expectation** and the **covariance** matrix for these distributions in order for the **posterior** distribution to be fully characterized.

The Kalman Filter

- For the **likelihood** the expectation is obtained by

$$\mathbb{E}\left[\mathbf{y}_n|\mathbf{x}_n\right] = \mathbb{E}\left[\mathbf{H}_n\mathbf{x}_n + \mathbf{v}_n\right] \qquad \text{using} \quad \mathbf{y}_n = \mathbf{H}_n\mathbf{x}_n + \mathbf{v}_n$$

Non-biased

$$= \mathbf{H}_n\mathbf{x}_n$$

- The covariance matrix is obtained by:

$$\operatorname{cov}\left[\mathbf{y}_n|\mathbf{x}_n\right] = \operatorname{cov}\left[\mathbf{v}_n\right]$$

$$= \mathbf{R}_n$$

The Kalman Filter

- Derivation of the equations
  - Thus the pdf of the **likelihood** can be written as

$$\pi(\mathbf{y}_n|\mathbf{x}_n) = \frac{1}{(2\pi)^{N_y/2}|\mathbf{R}_n|^{1/2}} \exp\left[-\frac{1}{2}\left(\mathbf{y}_n - \mathbf{H}_n\mathbf{x}_n\right)^T \mathbf{R}_n^{-1}\left(\mathbf{y}_n - \mathbf{H}_n\mathbf{x}_n\right)\right]$$

The Kalman Filter

- Derivation of the equations
  - For the **prior** pdf $\pi(\mathbf{x}_n|\mathbf{y}_{0:n-1})$, the expectation is obtained by:

$$\mathbb{E}\left[\mathbf{x}_n|\mathbf{y}_{0:n-1}\right] = \mathbb{E}\left[\mathbf{F}_n\widehat{\mathbf{x}}_{n-1} + \mathbf{w}_{n-1}\right]$$

$$= \mathbf{F}_n\widehat{\mathbf{x}}_{n-1}$$

$\updownarrow$ Equivalent

$$= \widehat{\mathbf{x}}_{n|n-1}$$

  - And the covariance matrix is obtained by:

Both using

$$\mathbf{x}_n = \mathbf{F}_n\mathbf{x}_{n-1} + \mathbf{w}_{n-1}$$

$$\operatorname{cov}\left[\mathbf{x}_n|\mathbf{y}_{0:n-1}\right] = \operatorname{cov}\left[\mathbf{x}_n - \widehat{\mathbf{x}}_{n|n-1}\right]$$

$$= \operatorname{cov}\left[\mathbf{F}_n\mathbf{x}_{n-1} + \mathbf{w}_{n-1} - \mathbf{F}_n\widehat{\mathbf{x}}_{n-1}\right]$$

$$= \operatorname{cov}\left[\mathbf{F}_n\left(\mathbf{x}_{n-1} - \widehat{\mathbf{x}}_{n-1}\right) + \mathbf{w}_{k-1}\right]$$

$$\operatorname{cov}[\mathbf{x}] = \mathrm{E}\left(\mathbf{x}\mathbf{x}^{\mathrm{T}}\right)$$

$$= \mathbf{F}_n\mathbf{P}_{n-1}\mathbf{F}_n^T + \mathbf{Q}_n$$

$\updownarrow$ Equivalent

$$= \mathbf{P}_{n|n-1}$$

The Kalman Filter

- Derivation of the equations
  - Thus the pdf of the **priori** can be written as

$$\pi(\mathbf{x}_n|\mathbf{y}_{0:n-1}) = \frac{1}{(2\pi)^{N_x/2}|\mathbf{P}_{n|n-1}|^{1/2}} \exp\left[-\frac{1}{2}\left(\mathbf{x}_n - \widehat{\mathbf{x}}_{n|n-1}\right)^T \mathbf{P}_{n|n-1}^{-1}\left(\mathbf{x}_n - \widehat{\mathbf{x}}_{n|n-1}\right)\right]$$

The Kalman Filter

- Derivation of the equations
  - The pdf of the **posterior** distribution is obtained combining both **prior** and **likelihood** pdf´s:

$$\pi\left(\mathbf{x}_n|\mathbf{y}_{0:n}\right) \propto \exp\left\{-\frac{1}{2}\left[\left(\mathbf{y}_n - \mathbf{H}_n\mathbf{x}_n\right)^T \mathbf{R}_n^{-1}\left(\mathbf{y}_n - \mathbf{H}_n\mathbf{x}_n\right) + \right.\right.$$
$$\left.\left. + \left(\mathbf{x}_n - \widehat{\mathbf{x}}_{n|n-1}\right)^T \mathbf{P}_{n|n-1}^{-1}\left(\mathbf{x}_n - \widehat{\mathbf{x}}_{n|n-1}\right)\right]\right\}$$

  - The logarithm of the posterior pdf is taken and the derivative is calculated as

$$\frac{\partial}{\partial \mathbf{x}_n}\left[\ln \pi(\mathbf{x}_n|\mathbf{y}_{0:n})\right] = \mathbf{H}_n^T\mathbf{R}_n^{-1}\left(\mathbf{y}_n - \mathbf{H}_n\mathbf{x}_n\right) - \mathbf{P}_{n|n-1}^{-1}\left(\mathbf{x}_n - \widehat{\mathbf{x}}_{n|n-1}\right)$$

The Kalman Filter

- Derivation of the equations
  - Taking x_n as $\hat{x}_n^{MAP}$ one obtains:

$$\boxed{\mathrm{IV}} = \boxed{\mathrm{I}}\boxed{\mathrm{III}}$$

$$\widehat{\mathbf{x}}_n^{\mathrm{MAP}} = \overbrace{\left(\mathbf{H}_n^T\mathbf{R}_n^{-1}\mathbf{H}_n + \mathbf{P}_{n|n-1}^{-1}\right)^{-1}}^{\boxed{\mathrm{I}}}\left(\overbrace{\mathbf{P}_{n|n-1}^{-1}}^{\boxed{\mathrm{II}}}\widehat{\mathbf{x}}_{n|n-1} + \overbrace{\mathbf{H}_n^T\mathbf{R}_n^{-1}}^{\boxed{\mathrm{III}}}\mathbf{y}_n\right)$$

  - By using the following lemma for the first term in the right hand side (I),

$$\left(\mathbf{A} + \mathbf{C}\mathbf{B}\mathbf{C}^T\right)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}\left(\mathbf{B}^{-1} + \mathbf{C}^T\mathbf{A}^{-1}\mathbf{C}\right)\mathbf{C}^T\mathbf{A}^{-1}$$

  - Where

$$\mathbf{A} = \mathbf{P}_{n|n-1}^{-1}, \quad \mathbf{B} = \mathbf{R}^{-1}, \quad \mathbf{C} = \mathbf{H}_n^T$$

  - One obtains,

$$\left(\mathbf{H}_n^T\mathbf{R}_n^{-1}\mathbf{H}_n + \mathbf{P}_{n|n-1}^{-1}\right)^{-1} = \mathbf{P}_{n|n-1} -$$
$$- \mathbf{P}_{n|n-1}\mathbf{H}_n^T\left(\mathbf{H}_n\mathbf{P}_{n|n-1}\mathbf{H}_n^T + \mathbf{R}_n\right)^{-1}\mathbf{H}_n\mathbf{P}_{n|n-1}$$

The Kalman Filter

- Derivation of the equations
  - A further simplification is obtained with the following lemma:

$$\left(\mathrm{P}^{-1} + \mathrm{B}^T\mathrm{R}^{-1}\mathrm{B}\right)^{-1}\mathrm{B}^T\mathrm{R}^{-1} = \mathrm{P}\mathrm{B}^T\left(\mathrm{B}\mathrm{P}\mathrm{B}^T + \mathrm{R}\right)^{-1} \qquad \boxed{\mathrm{IV}} = \boxed{\mathrm{I}}\,\boxed{\mathrm{III}}$$

  - Where

$$\mathrm{P}^{-1} = \mathrm{P}_{n|n-1}^{-1}, \quad \mathrm{B} = \mathrm{H}_n, \quad \mathrm{R} = \mathrm{R}_n$$

  - Then

$$\left(\mathrm{H}_n^T\mathrm{R}_n^{-1}\mathrm{H}_n + \mathrm{P}_{n|n-1}^{-1}\right)^{-1}\mathrm{H}_n^T\mathrm{R}_n^{-1} = \mathrm{P}_{n|n-1}\mathrm{H}_n^T\left(\mathrm{H}_n\mathrm{P}_{n|n-1}\mathrm{H}_n^T + \mathrm{R}_n\right)^{-1}$$

The Kalman Filter

- Derivation of the equations
  - Reminding: Derivative of $\ln(\pi(x|y))$ with respect to x

$$\hat{\mathbf{x}}_n^{\mathrm{MAP}} = \left(\mathbf{H}_n^T\mathbf{R}_n^{-1}\mathbf{H}_n + \mathbf{P}_{n|n-1}^{-1}\right)^{-1}\left(\mathbf{P}_{n|n-1}^{-1}\hat{\mathbf{x}}_{n|n-1} + \mathbf{H}_n^T\mathbf{R}_n^{-1}\mathbf{y}_n\right)$$

  - With the information of the previous lemmas:

$$\left(\mathbf{H}_n^T\mathbf{R}_n^{-1}\mathbf{H}_n + \mathbf{P}_{n|n-1}^{-1}\right)^{-1} = \mathbf{P}_{n|n-1} - \\ - \mathbf{P}_{n|n-1}\mathbf{H}_n^T\left(\mathbf{H}_n\mathbf{P}_{n|n-1}\mathbf{H}_n^T + \mathbf{R}_n\right)^{-1}\mathbf{H}_n\mathbf{P}_{n|n-1}$$

$$\left(\mathbf{H}_n^T\mathbf{R}_n^{-1}\mathbf{H}_n + \mathbf{P}_{n|n-1}^{-1}\right)^{-1}\mathbf{H}_n^T\mathbf{R}_n^{-1} = \mathbf{P}_{n|n-1}\mathbf{H}_n^T\left(\mathbf{H}_n\mathbf{P}_{n|n-1}\mathbf{H}_n^T + \mathbf{R}_n\right)^{-1}$$

  - Results in $\hat{\mathbf{x}}_n^{\mathrm{MAP}} = \hat{\mathbf{x}}_{n|n-1} + \mathbf{K}_n\left(\mathbf{y}_n - \mathbf{H}_n\hat{\mathbf{x}}_{n|n-1}\right)$
  - Where $\mathbf{K}_n = \mathbf{P}_{n|n-1}\mathbf{H}_n^T\left(\mathbf{H}_n\mathbf{P}_{n|n-1}\mathbf{H}_n^T + \mathbf{R}\right)^{-1}$

The Kalman Filter

- Derivation of the equations
  - Covariance matrix derivation

$$
\begin{aligned}
\mathbf{P}_n &= \operatorname{cov}\left[\mathbf{x}_n - \widehat{\mathbf{x}}_n^{\mathrm{MAP}}\right] \\
&= \operatorname{cov}\left[\mathbf{x}_n - \widehat{\mathbf{x}}_{n|n-1} - \mathbf{K}_\mathbf{n}\left(\mathbf{y}_n - \mathbf{H}_n\widehat{\mathbf{x}}_{n|n-1}\right)\right] \\
&= \operatorname{cov}\left[\left(\mathbf{I} - \mathbf{K}_n\mathbf{H}_n\right)\left(\mathbf{x}_n - \widehat{\mathbf{x}}_{n|n-1}\right) + \mathbf{K}_\mathbf{n}\mathbf{v}_n\right] \\
&= \left(\mathbf{I} - \mathbf{K}_n\mathbf{H}_n\right)\mathbf{P}_{n|n-1}\left(\mathbf{I} - \mathbf{K}_n\mathbf{H}_n\right)^T + \mathbf{K}_n\mathbf{R}_n\mathbf{K}_n^T
\end{aligned}
$$

$$\mathbf{K}_n\mathbf{H}_n\left(\mathbf{x}_n - \mathbf{x}_n\right)$$

$$\operatorname{cov}[\mathbf{x}] = \mathrm{E}\left(\mathbf{x}\mathbf{x}^{\mathrm{T}}\right)$$

  - By using

$$\mathbf{K}_n = \mathbf{P}_{n|n-1}\mathbf{H}_n^T\left(\mathbf{H}_n\mathbf{P}_{n|n-1}\mathbf{H}_n^T + \mathbf{R}\right)^{-1}$$

  - One obtain

$$\mathbf{P}_n = \left(\mathbf{I} - \mathbf{K}_n\mathbf{H}_n\right)\mathbf{P}_{n|n-1}$$

The Kalman Filter

Prediction:

$$\hat{\mathbf{x}}_k^- = \mathbf{F}_k \hat{\mathbf{x}}_{k-1}^+$$

$$\mathbf{P}_k^- = \mathbf{F}_k \mathbf{P}_{k-1}^+ \mathbf{F}_k^T + \mathbf{Q}_k$$

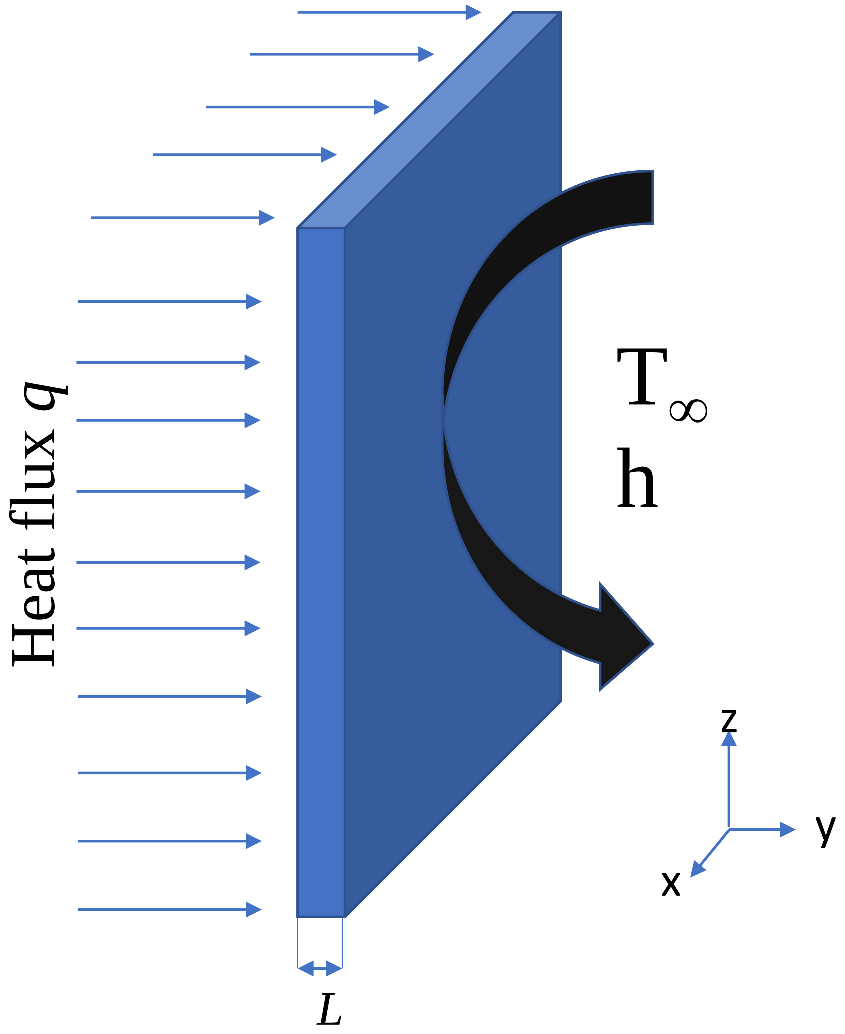
Update:

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1}$$

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \underbrace{\left(\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^- \right)}_{\text{Measurement Innovation}}$$

$$\mathbf{P}_k^+ = \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k \right) \mathbf{P}_k^-$$

Example I: Lumped System



$$\begin{cases} \dfrac{d\theta}{dt} + m\theta(t) = \dfrac{mq(t)}{h} & \text{For } t > 0 \\ \theta(0) = \theta_o & \text{For } t = 0 \\ \text{where} & \\ \theta = T(t) - T_\infty & m = \dfrac{h}{\rho c L} \\ \theta_o = T_o - T_\infty & \end{cases}$$

Example I: Lumped System

- Two illustrative cases are examined:
  - Heat flux $q(t)=q_o$ constant and deterministically known;
  - Heat flux $q(t)=q_o f(t)$ with unknown time variation;
- Plate is made of **aluminum** (ρ = 2707 kgm^{-3}, c = 896 $Jkg^{-1}K^{-1}$), with thickness L = 0.03 m, q_0 = 8000 W/m^2, T_∞=20 ^oC, h = 50 $Wm^{-2}K^{-1}$ and T_0 = 50 ^oC.
- **Measurement** of the transient temperature of the slab are **assumed available**. These **measurements** contain **additive, uncorrelated, Gaussian error**, with zero mean and a constant standard deviation σ_z
- The **errors in the state evolution model** are also supposed to be additive, uncorrelated, Gaussian, with zero mean and a constant standard deviation σ_θ

Example I: Lumped System

- (i) Heat Flux $q(t)=q_o$ constant and deterministically known
  - The analytical solution for this problem is given by:

$$\theta(t) = \theta_o e^{-mt} + \frac{q_o}{h}\left(1 - e^{-mt}\right)$$

  - The only state variable in this case is the temperature $\theta(t_k)= \theta_k$ since the applied heat flux q_o is constant and deterministically known, as the other parameter appearing in the formulation. By using a forward finite differences approximation for the time derivative in equation:

$$\frac{d\theta}{dt} + m\theta(t) = \frac{mq(t)}{h}$$

We obtain:

$$\theta_k = \left(1 - m\Delta t\right)\theta_{k-1} + \frac{mq_o}{h}\Delta t$$

Example I: Lumped System

- Therefore, the state and observation models given by:

$$\boxed{\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{s}_{k-1} + \mathbf{w}_{k-1}} \qquad \boxed{\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{n}_k}$$

are obtained with

$$\mathbf{x}_k = \left[\theta_k\right] \qquad \mathbf{s}_k = \left[m\frac{q_o}{h}\Delta t\right] \qquad \mathbf{Q}_k = \left[\sigma_\theta^2\right]$$

$$\mathbf{F}_k = \left[\left(1 - m\Delta t\right)\right] \qquad \mathbf{H}_k = \left[1\right] \qquad \mathbf{R}_k = \left[\sigma_z^2\right]$$

Example I: Lumped System

- (ii) Heat Flux $q(t)=q_0\, f(t)$ with unknown time variation,
  - The analytical solution for this problem is given by:

$$\theta(t)=\left\{\theta_o+\frac{mq_o}{h}\int_{t'=0}^{t} e^{mt'} f(t')dt'\right\}$$

  - In this case, **the state variables** are given by the temperatures $\boldsymbol{\theta(t_k)=\theta_k}$ and the function that gives the time variation of the applied heat flux, that is $\boldsymbol{f(t_k)=f_k}$. As in the case examined previously, the applied heat flux $\mathbf{q_o}$ is **constant and deterministically known**, as the **other parameters appearing in the formulation**. By using a forward **finite-difference approximation** for the time derivative we obtain the equation for the evolution of the state variable $\theta(t_k)=\theta_k$:

$$\theta_k=(1-m\Delta t)\theta_{k-1}+\left(\frac{mq_o}{h}\Delta t\right)f_{k-1}$$

  - A random walk model is used for the state variable $f(t_k)=f_{k,}$, which is given in the form

$$f_k=f_{k-1}+\varepsilon_{k-1}$$

  - Where ε_{k-1} is Gaussian with zero mean and constant standard deviation σ_{rw}

Example I: Lumped System

- (ii) Heat Flux $q(t) = q_o f(t)$ with unknown time variation
  - Therefore, the state and observation models given by:

$$\boxed{\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{s}_{k-1} + \mathbf{w}_{k-1}}$$

$$\boxed{\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{n}_k}$$

are obtained with

$$\mathbf{s}_k = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mathbf{Q}_k = \begin{bmatrix} \sigma_\theta^2 & 0 \\ 0 & \sigma_{rw}^2 \end{bmatrix}$$

$$\mathbf{F}_k = \begin{bmatrix} (1 - m\Delta t) & \dfrac{mq_o}{h}\Delta t \\ 0 & 1 \end{bmatrix} \qquad \mathbf{x}_k = \begin{bmatrix} \theta_k \\ f_k \end{bmatrix}$$

$$\mathbf{R}_k = \begin{bmatrix} \sigma_z^2 \end{bmatrix}$$

$$\mathbf{H}_k = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

Example II: Experimental result

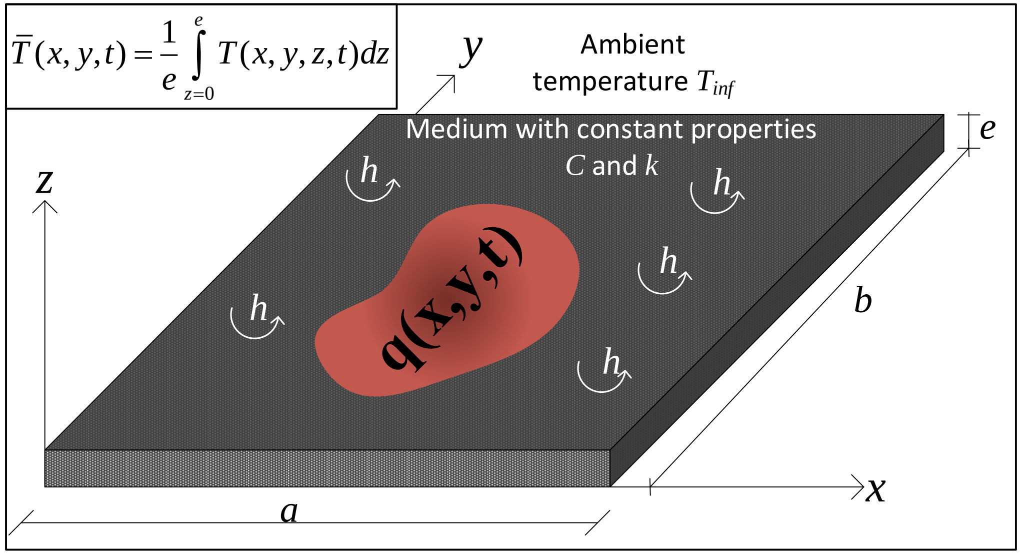
- Estimation of position-dependent transient heat source

$$C\frac{\partial \bar{T}}{\partial t}=\frac{\partial}{\partial x}\left(k\frac{\partial \bar{T}}{\partial x}\right)+\frac{\partial}{\partial y}\left(k\frac{\partial \bar{T}}{\partial y}\right)-\frac{h}{e}\left(\bar{T}-T_\infty\right)+\frac{q(x,y,t)}{e} \quad \text{in } 0<x<\text{a},\ 0<y<\text{b, for } t>0$$

$$\frac{\partial \bar{T}}{\partial y}=0 \quad \text{at } x=0 \text{ and } x=\text{a, for } t>0$$

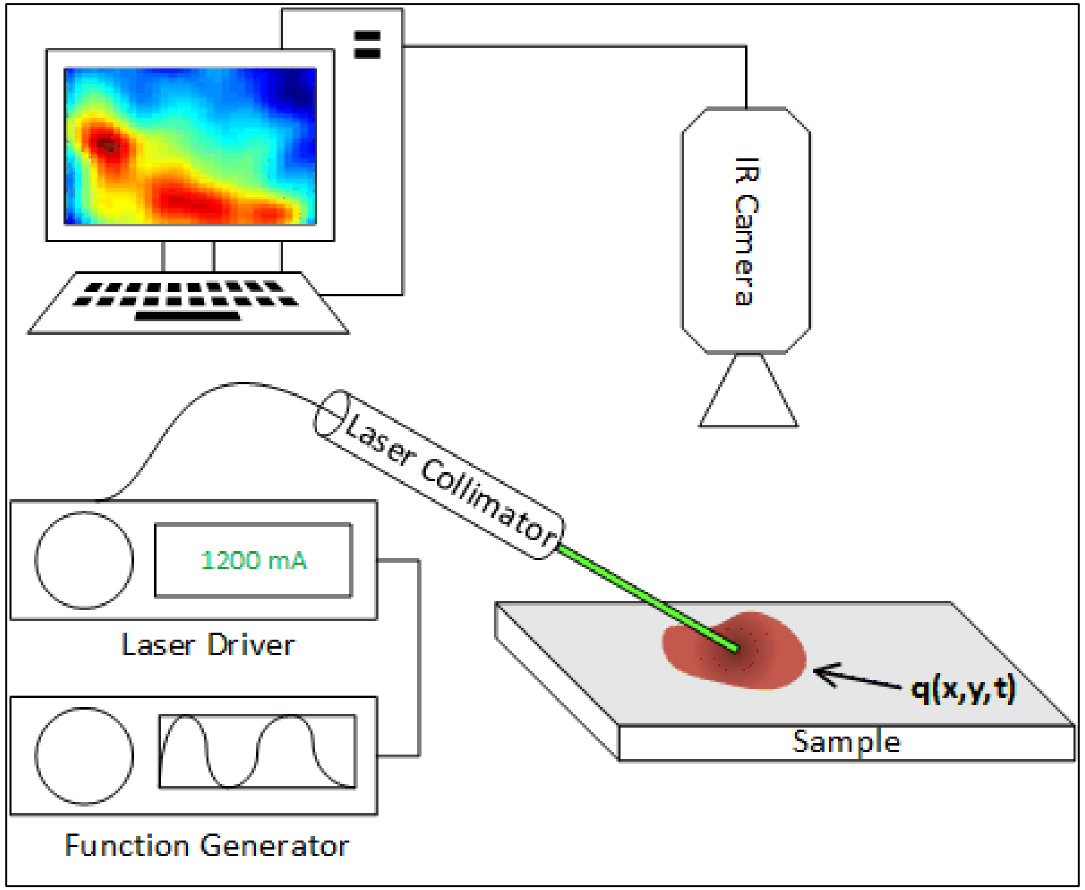
$$\frac{\partial \bar{T}}{\partial x}=0 \quad \text{at } y=0 \text{ and } y=\text{b, for } t>0$$

$$\bar{T}=T_0 \quad \text{for } t=0,\ \text{in } 0<x<\text{a and } 0<y<\text{b}$$



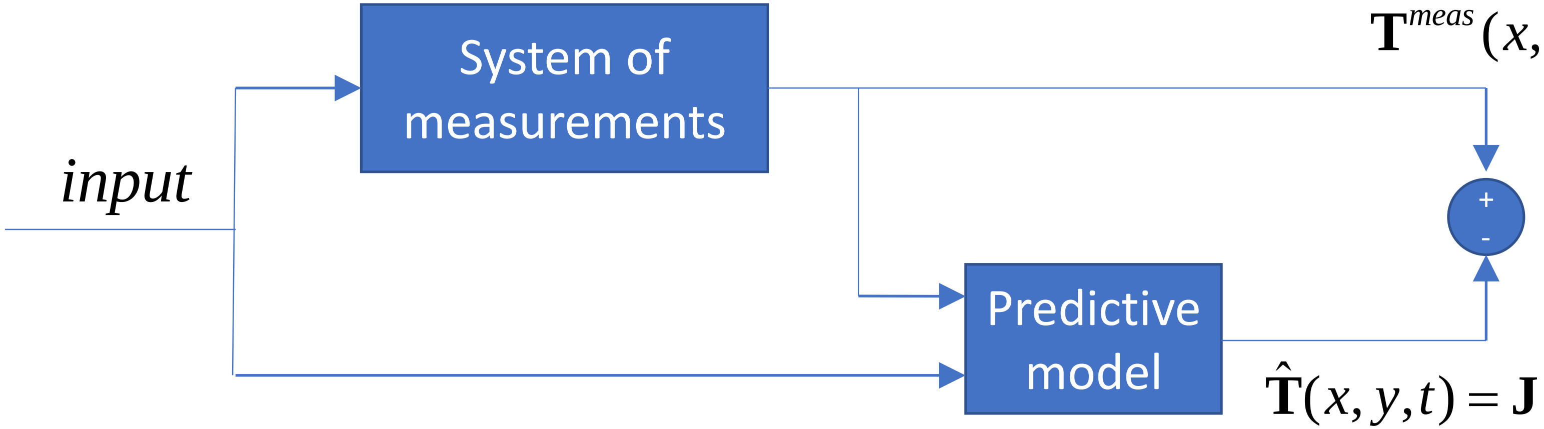
Example II: Experimental result

- Schematics of the experiment



Example II: Experimental result

- Nodal approach



$$\mathbf{T}^{meas}(x, y, t)$$

$$\mathbf{P}^{\mathrm{OLS}} = \left(\mathbf{J}^{\mathrm{T}}\mathbf{J}\right)^{-1}\mathbf{J}^{\mathrm{T}}\hat{\mathbf{T}}$$

$$\hat{\mathbf{T}}(x, y, t) = \mathbf{J}\left(\mathbf{T}^{meas}\right)\mathbf{P}$$

$$\frac{\partial T}{\partial t} = \alpha(x, y)\nabla^2 T - H(x, y)\left(T - T_\infty\right) + G(x, y, t)$$

$$Y_{i,j}^{n+1} = \Delta t\left(\frac{T_{i-1,j}^{n} - 2T_{i,j}^{n} + T_{i+1,j}^{n}}{(\Delta x)^2} + \frac{T_{i,j-1}^{n} - 2T_{i,j}^{n} + T_{i,j+1}^{n}}{(\Delta y)^2}\right)\alpha_{i,j} - \Delta t(T_{i,j}^{n} - T_\infty)H_{i,j} + \Delta t G_{i,j}$$

Example II: Experimental result

- Nodal approach for KF estimation

$$\mathbf{T}^{k+1} = \mathbf{T}^{k} + \mathbf{J}^{k}\mathbf{P}^{k} + \mathbf{v}_{T}^{k+1} \qquad \mathbf{P}^{k+1} = \mathbf{I}\mathbf{P}^{k} + \mathbf{v}_{P}^{k}$$

$$\mathbf{P}^{k} = \begin{bmatrix} \begin{bmatrix} a_1^k \\ H_1^k \\ G_1^k \end{bmatrix} \\ \begin{bmatrix} a_2^k \\ H_2^k \\ G_2^k \end{bmatrix} \\ \vdots \\ \begin{bmatrix} a_M^k \\ H_M^k \\ G_M^k \end{bmatrix} \end{bmatrix} \qquad \mathbf{T}^{k} = \begin{bmatrix} T_1^k \\ T_2^k \\ \vdots \\ T_M^k \end{bmatrix}$$

$$\mathbf{J}^{k} = \begin{bmatrix} L_1^k & -\Delta t(T_1^k - T_\infty) & \Delta t & 0 & 0 & & & \\ 0 & 0 & 0 & L_2^k & -\Delta t(T_2^k - T_\infty) & \Delta t\, 0 & & \\ \vdots & & & & & & & \\ 0 & \cdots & & & & \cdots 0\, L_M^k & -\Delta t(T_M^k - T_\infty) & \Delta t \end{bmatrix}$$

Example II: Experimental result

- Therefore, the state and observation models given by:

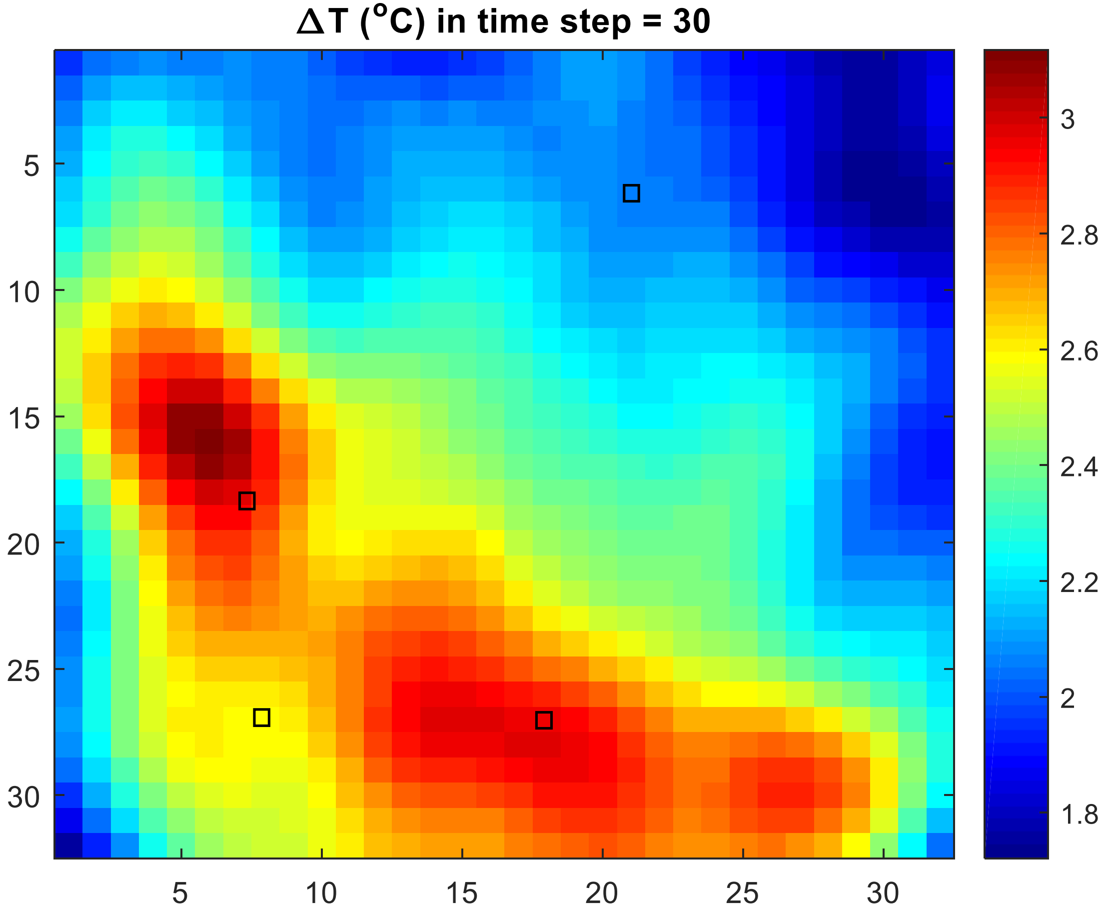
$$\boxed{\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{s}_{k-1} + \mathbf{w}_{k-1}} \qquad \boxed{\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{n}_k}$$

are obtained with

$$\mathbf{x}_k = \begin{bmatrix} \mathbf{T}_k \\ \mathbf{P}_k \end{bmatrix} \qquad \mathbf{F}^k = \begin{bmatrix} \mathbf{I} & \mathbf{J}^k \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

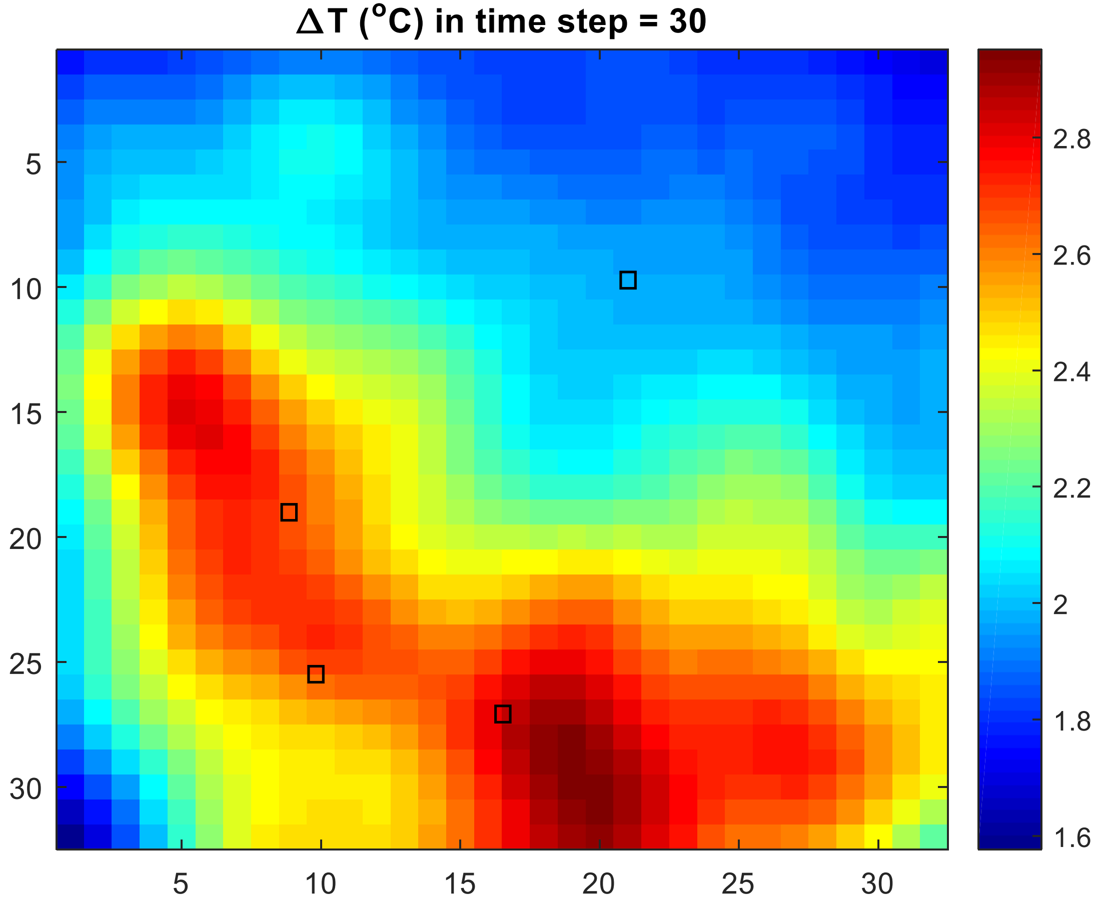
$$\mathbf{H}^k = [\mathbf{I} \;\; \mathbf{0}] \qquad \mathbf{Q}_k = \begin{bmatrix} \mathbf{Q}_T^k & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_P^k \end{bmatrix}$$

Example II: Experimental result

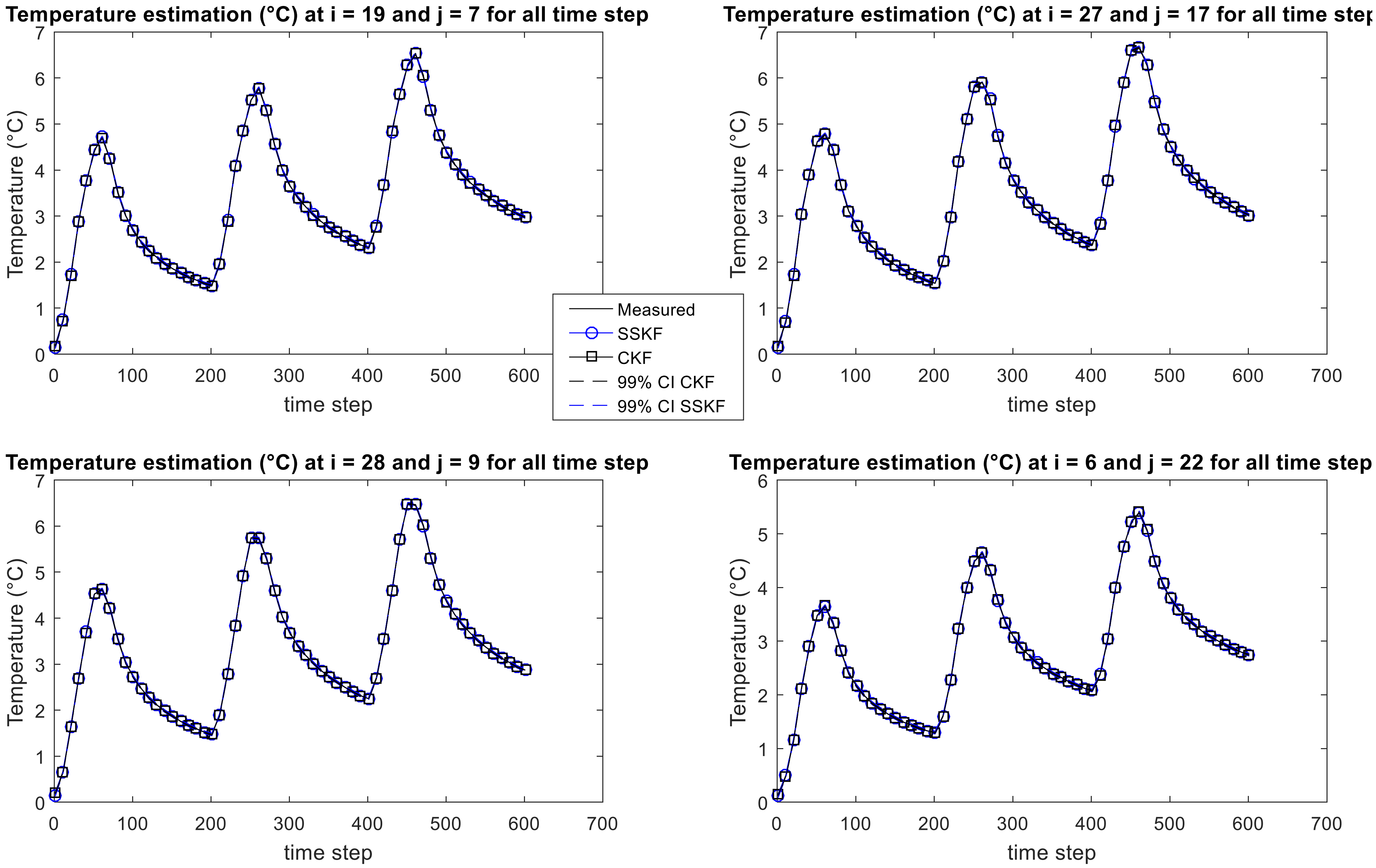


Sinusoidal pulse

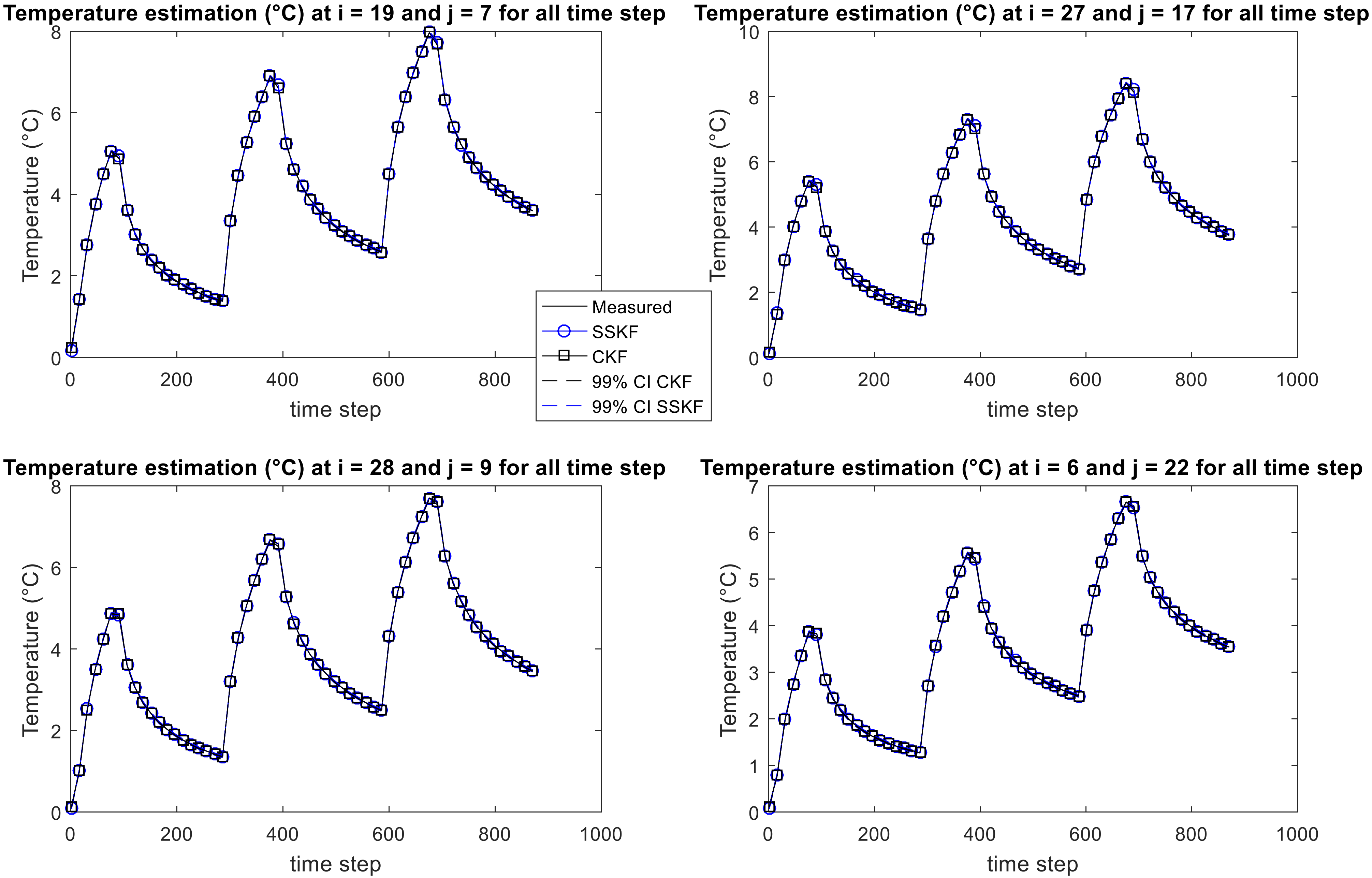
Square pulse



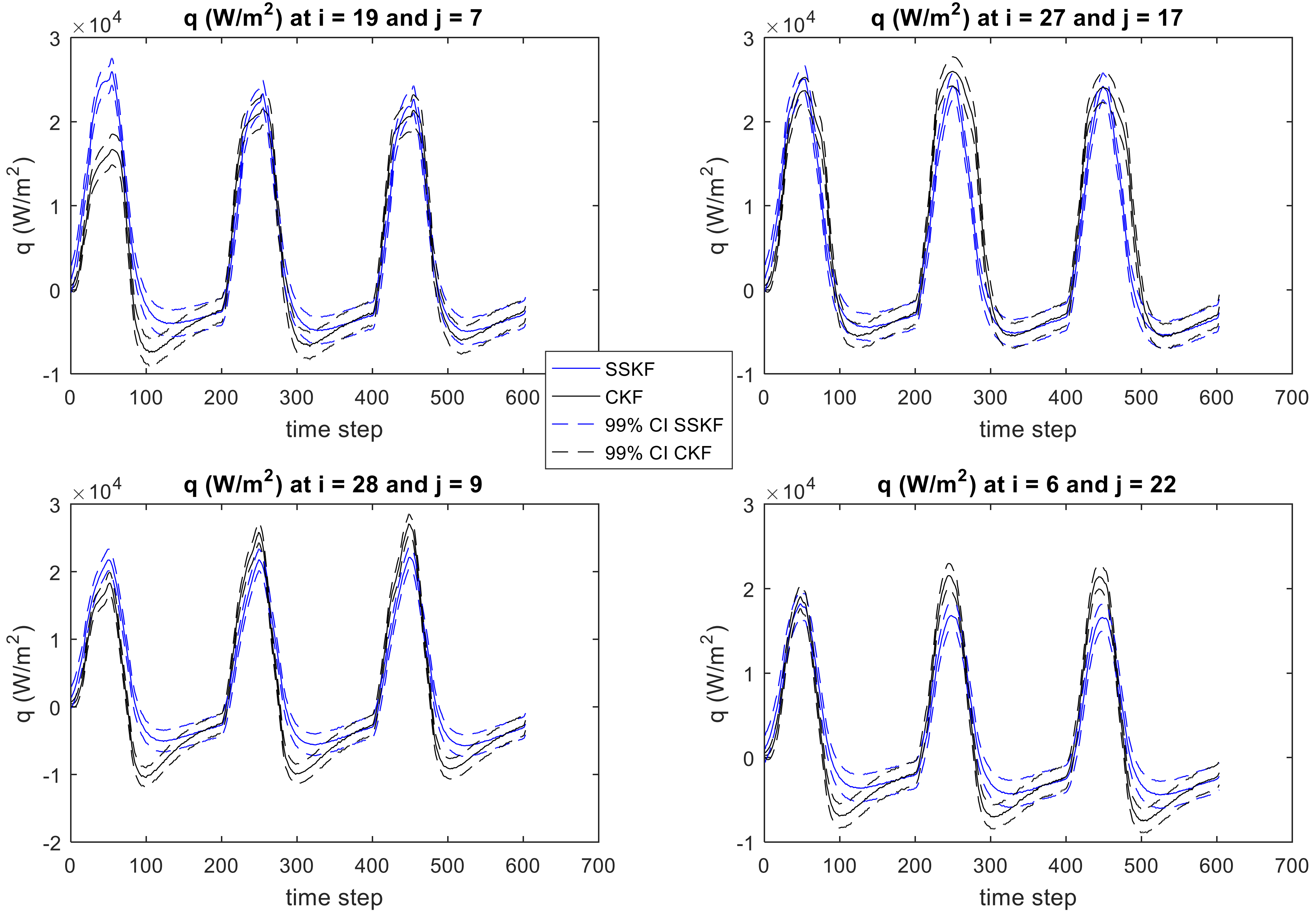
Example II: Experimental result - Sine



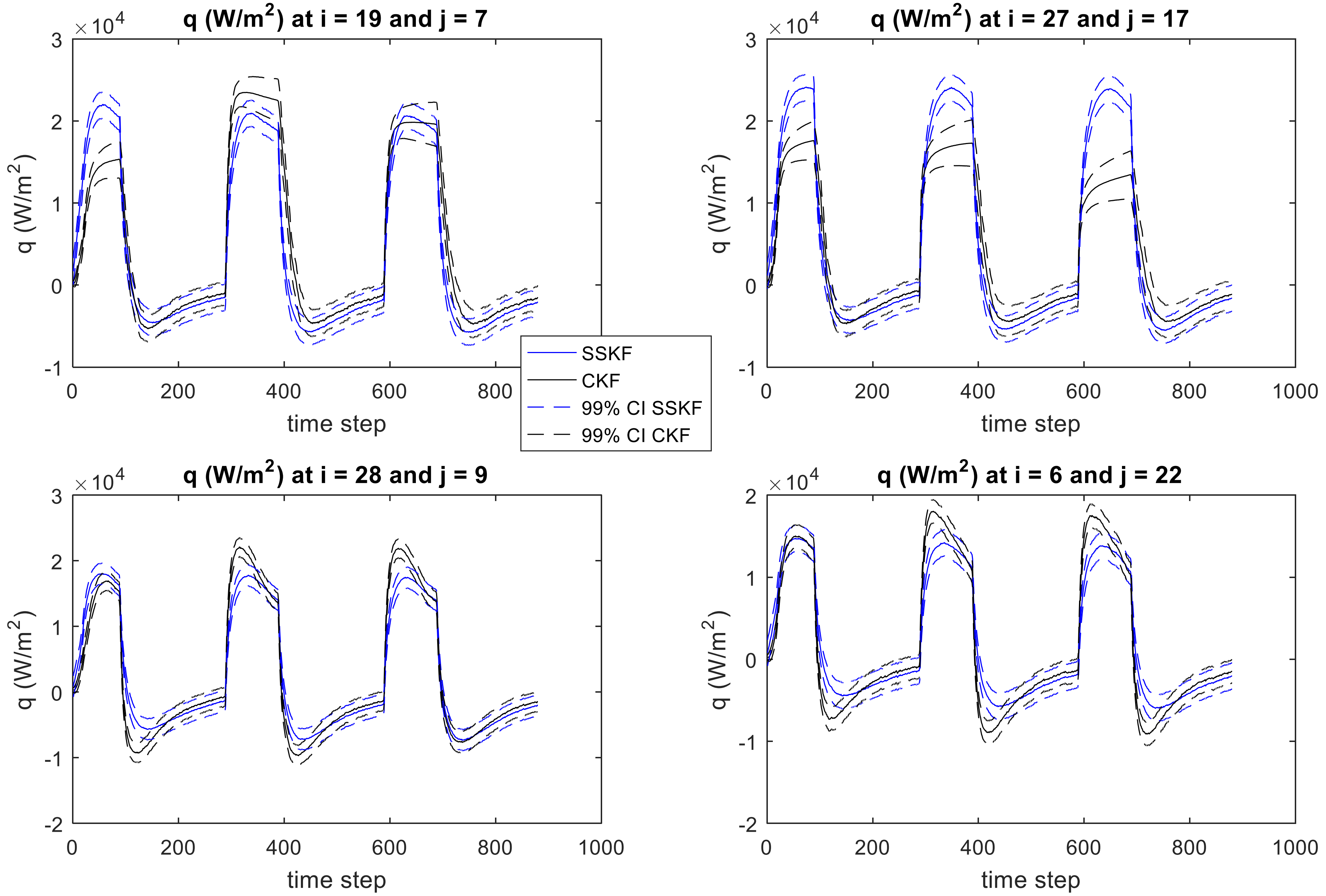
Example II: Experimental result - Square



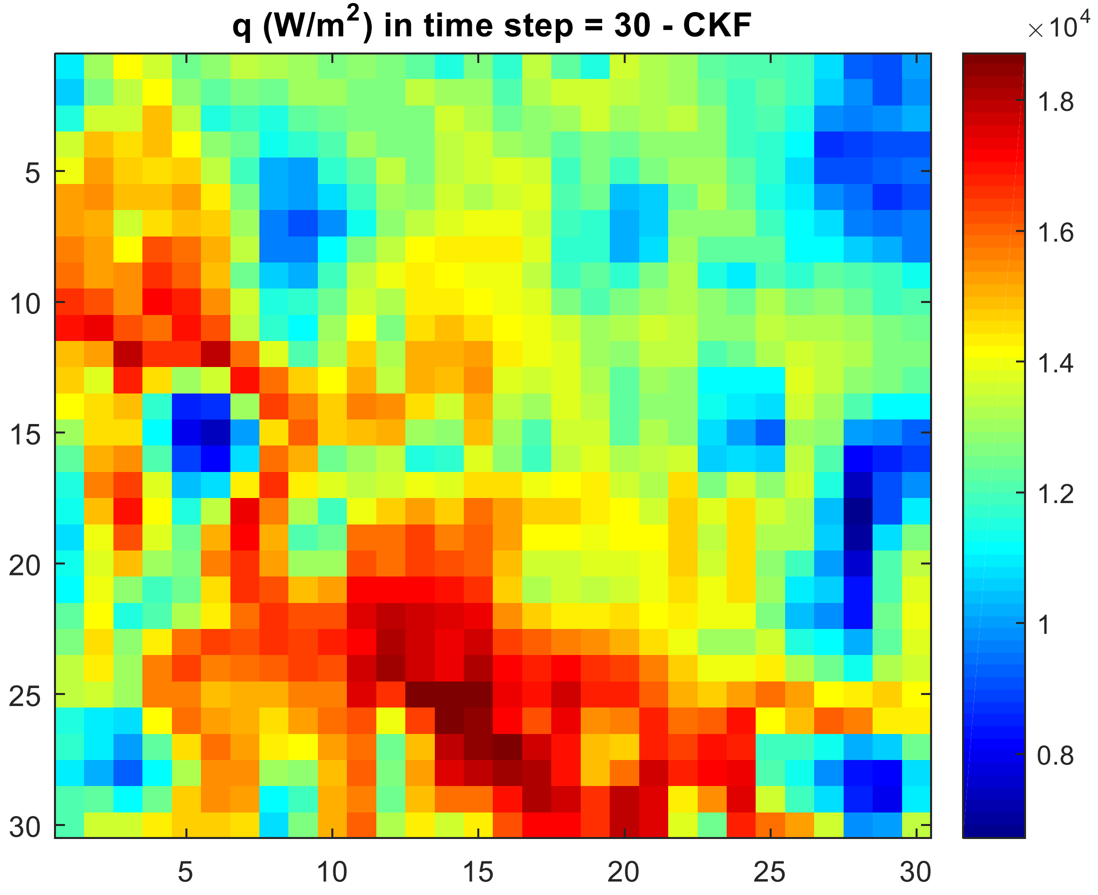
Example II: Experimental result - Sine



Example II: Experimental result - Square

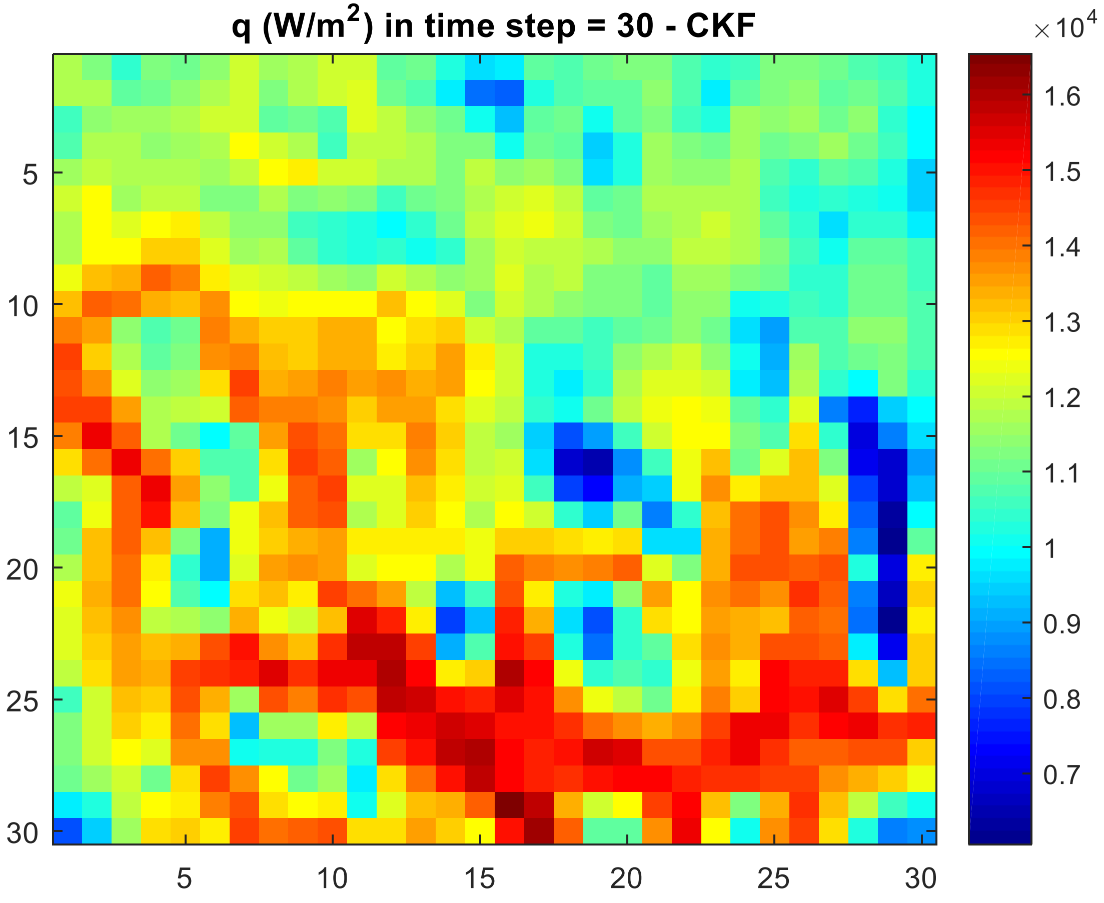


Example II: Experimental result



Sin

Square



The Kalman Filter

- References
  - WELCH, G., E BISHOP, G., “An Introduction to the Kalman Filter”, (2006). Available at http://www.cs.unc.edu/~welch/media/pdf/kalman_intro.pdf
  - SORENSON, H. W., “Least-squares estimation: from Gauss to Kalman”, *IEEE Spectrum* (1970), pp. 63-68.
  - MAYBECK, P., 1979, Stochastic models, estimation and control, *Academic Press*, New York.

The Kalman Filter

- Thank you!

The Kalman Filter – Extensions

Fact: Kalman filter is restrict to **linear** and **Gaussian** problems.

One investigation:

- System matrices are allowed to change with time.
- What if they don't?

Second investigation:

- What if the hypothesis above do not hold?
- Can we solve **nonlinear and/or non Gaussian** problems?

The Kalman Filter – Extensions

- What if instead of

$$\boxed{\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{u}_{k-1} + \mathbf{w}_{k-1}} \qquad \boxed{\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{n}_k}$$

we had its time-invariant version:

$$\boxed{\mathbf{x}_k = \mathbf{F} \mathbf{x}_{k-1} + \mathbf{u}_{k-1} + \mathbf{w}_{k-1}} \qquad \boxed{\mathbf{y}_k = \mathbf{H} \mathbf{x}_k + \mathbf{n}_k}$$

with the noise given by

$$\boxed{\mathbf{w}_k \sim N\left(\mathbf{0}, \mathbf{Q}\right)} \qquad \boxed{\mathbf{n}_k \sim N\left(\mathbf{0}, \mathbf{R}\right)}$$

The Kalman Filter – Extensions

In other words

$$\boxed{\mathbf{F}_k = \mathbf{F}, \quad \mathbf{H}_k = \mathbf{H}, \quad \mathbf{Q}_k = \mathbf{Q}, \quad \mathbf{R}_k = \mathbf{R}}$$

As a practical result, **K** and **P** matrices behave asymptotically.

After a number of steps:

$$\boxed{\mathbf{K}_k \simeq \mathbf{K}_\infty, \quad \mathbf{P}_k^- \simeq \mathbf{P}_k^+ \simeq \mathbf{P}_\infty}$$

Can we use this to our advantage?

Steady-State Kalman Filter

Applying this result to Kalman filter equations yields:

$$\boxed{\mathbf{P}_\infty = \mathbf{F}\mathbf{P}_\infty\mathbf{F}^T - \mathbf{F}\mathbf{P}_\infty\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_\infty\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{P}_\infty\mathbf{F}^T + \mathbf{Q}}$$

$$\boxed{\mathbf{K}_\infty = \mathbf{P}_\infty\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_\infty\mathbf{H}^T + \mathbf{R}\right)^{-1}}$$

$$\boxed{\hat{\mathbf{x}}_n^+ = \left(\mathbf{I} - \mathbf{K}_\infty\mathbf{H}\right)\mathbf{F}\hat{\mathbf{x}}_{n-1}^+ + \mathbf{K}_\infty\mathbf{y}_n}$$

These equations are referred to as the *Steady-State Kalman Filter (SSKF).*

Steady-State Kalman Filter

What are the properties of this method?

- Equations 1 and 2: *offline;*
- Equation 3: *online*;

Thus:

- **P** and **K** are calculated *offline*;
- **x** and **y** appear in one equation;
- No online matrix inversion;
- O(n²), instead of O(n³);

$$\boxed{\mathbf{P}_\infty = \mathbf{F}\mathbf{P}_\infty\mathbf{F}^T - \mathbf{F}\mathbf{P}_\infty\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_\infty\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{P}_\infty\mathbf{F}^T + \mathbf{Q}}$$

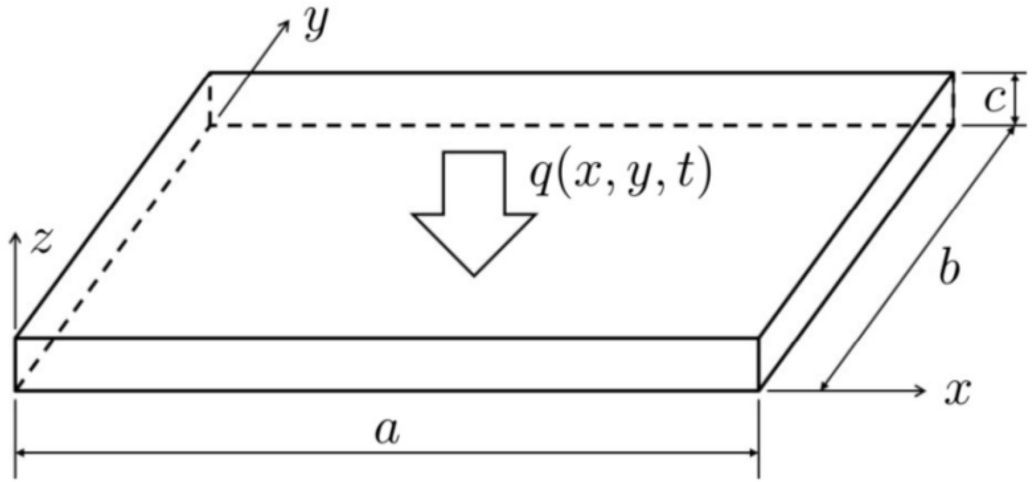
$$\boxed{\mathbf{K}_\infty = \mathbf{P}_\infty\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_\infty\mathbf{H}^T + \mathbf{R}\right)^{-1}}$$

$$\boxed{\hat{\mathbf{x}}_n^+ = \left(\mathbf{I} - \mathbf{K}_\infty\mathbf{H}\right)\mathbf{F}\hat{\mathbf{x}}_{n-1}^+ + \mathbf{K}_\infty\mathbf{y}_n}$$

Example III – Introduction

Heat Flux Identification problem:

- Heating of a flat square plate;
- 3D Nonlinear Heat Conduction;
- High magnitude heat flux;
- Measurements taken at opposite side;
- (a,b,c)=(120,120,30) mm



3D Nonlinear Transient Inverse Heat Conduction Problem:

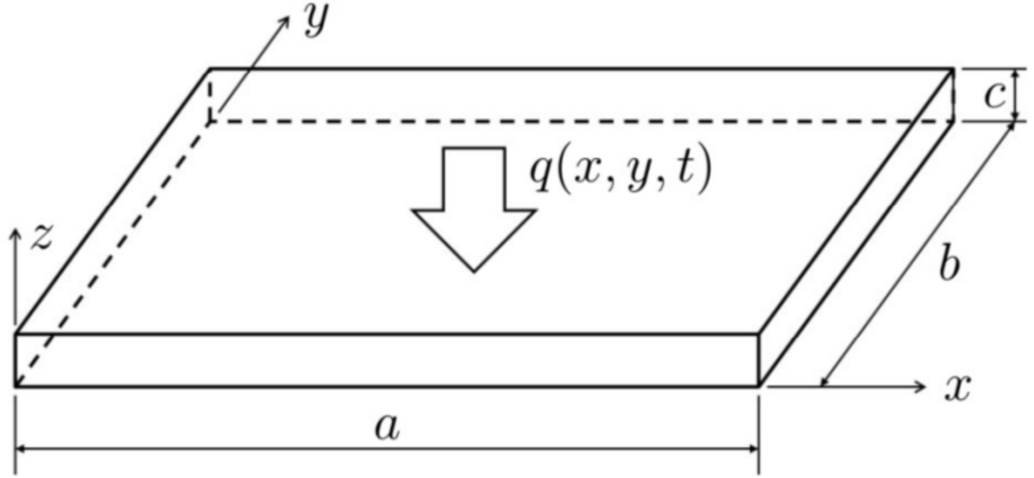
How to estimate the heat flux at real time?

Example III – Mathematical Modeling

Complete Mathematical Model:

- Governing Equation (t > 0):

$$C\left(T_c\right)\frac{\partial T_c}{\partial t} = \nabla \bullet \left[k_c\left(T_c\right)\nabla T_c\right],$$



- Initial Condition (t = 0):

$$T\left(x, y, z, t\right) = T_0,$$

Thermal Properties

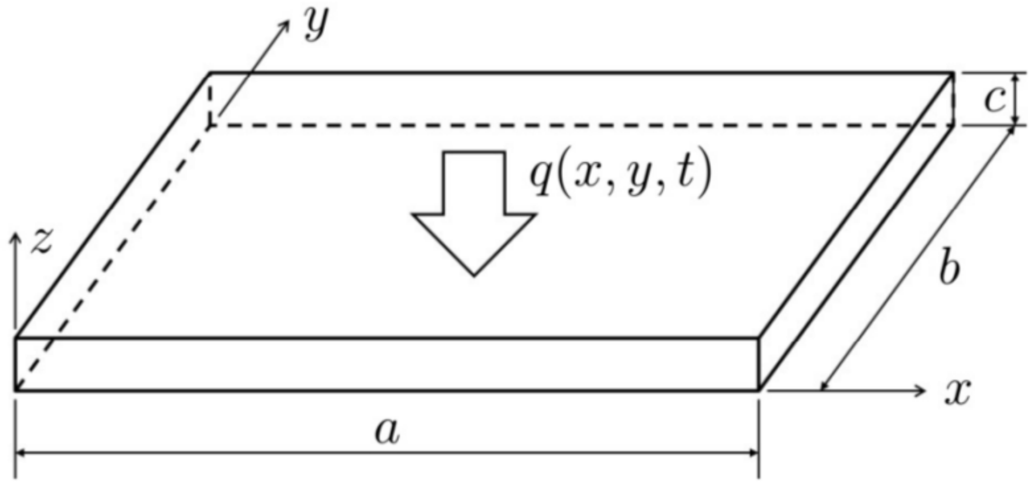
$$C\left(T\right) = 1324.75T + 3557900 \ [\mathrm{J/m^3}]$$

$$k\left(T\right) = 12.45 + 0.014T + 2.5171\times10^{-6}\mathrm{T}^2 \ [\mathrm{W/mK}]$$

Example III – Mathematical Modeling

Complete Mathematical Model:

- Boundary Conditions



$$\text{in } \; x=0 \text{ and } x=a: \quad \frac{\partial T_c}{\partial x}=0$$

$$\text{in } \; y=0 \text{ and } y=b: \quad \frac{\partial T_c}{\partial y}=0$$

$$\text{in } \; z=0: \quad \frac{\partial T_c}{\partial z}=0 \qquad\qquad \text{in } \; z=c: \quad k\left(T_c\right)\frac{\partial T_c}{\partial z}=q\left(x,y,t\right)$$

Example III – Mathematical Modeling

Complete Mathematical Model:

- Nonlinear: unable to apply Kalman filter for this case;
- 3D inverse problem might lead to high computational effort;

Proposal: Use a reduced model:

- Linearize it;
- Use mean temperature at the z-direction instead of actual temperature;
- Approximate temperature gradients across the thickness.

$$\boxed{\bar{T}\left(x,y,t\right)=\frac{1}{c}\int_{z=0}^{c}T(x,y,z,t)dz}$$

Example III – Mathematical Modeling

Reduced Mathematical Model:

- Governing Equation (t > 0):

$$C^* \frac{\partial \bar{T}}{\partial t} = k^* \frac{\partial^2 \bar{T}}{\partial x^2} + k^* \frac{\partial^2 \bar{T}}{\partial y^2} + \frac{q(x,y,t)}{c},$$

- Initial Condition (t = 0):

$$\bar{T}(x,y,t) = T_0,$$

Thermal Properties

$$C^* = C\left(T^*\right)$$

$$k^* = k\left(T^*\right)$$

w/ $T^* = 600$ K

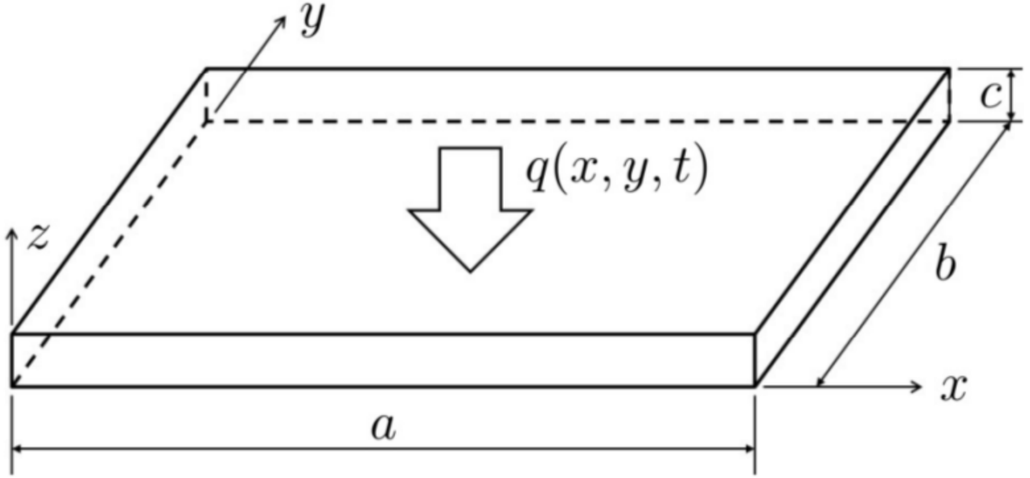
Example III – Mathematical Modeling

Reduced Mathematical Model:

- Boundary Conditions

$$\text{in } x=0 \text{ and } x=a: \quad \frac{\partial \bar{T}}{\partial x}=0$$

$$\text{in } y=0 \text{ and } y=b: \quad \frac{\partial \bar{T}}{\partial y}=0$$



Improved Lumped Method

- Temperature at z=0:

$$\boxed{T(x,y,0,t) \simeq \bar{T}(x,y,t) - \frac{c}{6k^*} q(x,y,t)}$$

Example III – Inverse Problem Settings

State vector:

- Mean temperature;
- Heat flux;
- Values throughout the mesh.

$$\mathbf{x}_k = \begin{bmatrix} \bar{\mathbf{T}}_k \\ \mathbf{q}_k \end{bmatrix}$$

Observation vector:

- Temperature at z=0;
- Values throughout the mesh.

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k \qquad \mathbf{H} = -\frac{c}{6k^*}\begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix}$$

Noise Covariance Matrices:

- Assumed uncorrelated.

$$\mathbf{Q} = \begin{bmatrix} \sigma_{\bar{T}}^2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_q^2\mathbf{I} \end{bmatrix} \qquad \mathbf{R} = \sigma_T^2\mathbf{I}$$

Example III – Inverse Problem Settings

Synthetic Measurements:

- Solution of forward complete problem;
- Achieved grid/time-step independence;
- Exact (Reference) heat flux;
- **<u>Total time: 2.0 s</u>** (200 measurements);

Inverse Problem:

- Reduced model;
- 24 x 24 grid with 0.02 s time step;
- No inverse crime;

$$\boxed{\mathbf{y}_n \sim \mathrm{N}\left(\mathbf{y}_n^{\text{exato}}, \mathbf{R}\right)}$$

$$\boxed{\mathbf{y}_n^{\text{exato}} = \mathbf{H}\mathbf{x}_n^{\text{exato}}}$$

$$\boxed{\mathbf{y}_n = \mathbf{y}_n^{\text{exato}} + \sigma_{\mathbf{y}}\boldsymbol{\omega}}$$

$$\text{OBS: } \boldsymbol{\omega} \sim \mathrm{N}\left(\mathbf{0}, \mathbf{I}\right)$$

Example III – Inverse Problem Settings

Exact (Reference) Heat Flux:

$$q(x,y,t)=\begin{cases} q_1, & \begin{matrix} x_{1,1}\le x\le x_{1,2}, \\ y_{1,1}\le y\le y_{1,2}, \end{matrix} \quad t\ge t_0 \\ q_2 & \begin{matrix} x_{2,1}\le x\le x_{2,2}, \\ y_{2,1}\le y\le y_{2,2}, \end{matrix} \quad t\ge t_0 \\ 0, & \text{otherwise} \end{cases}$$

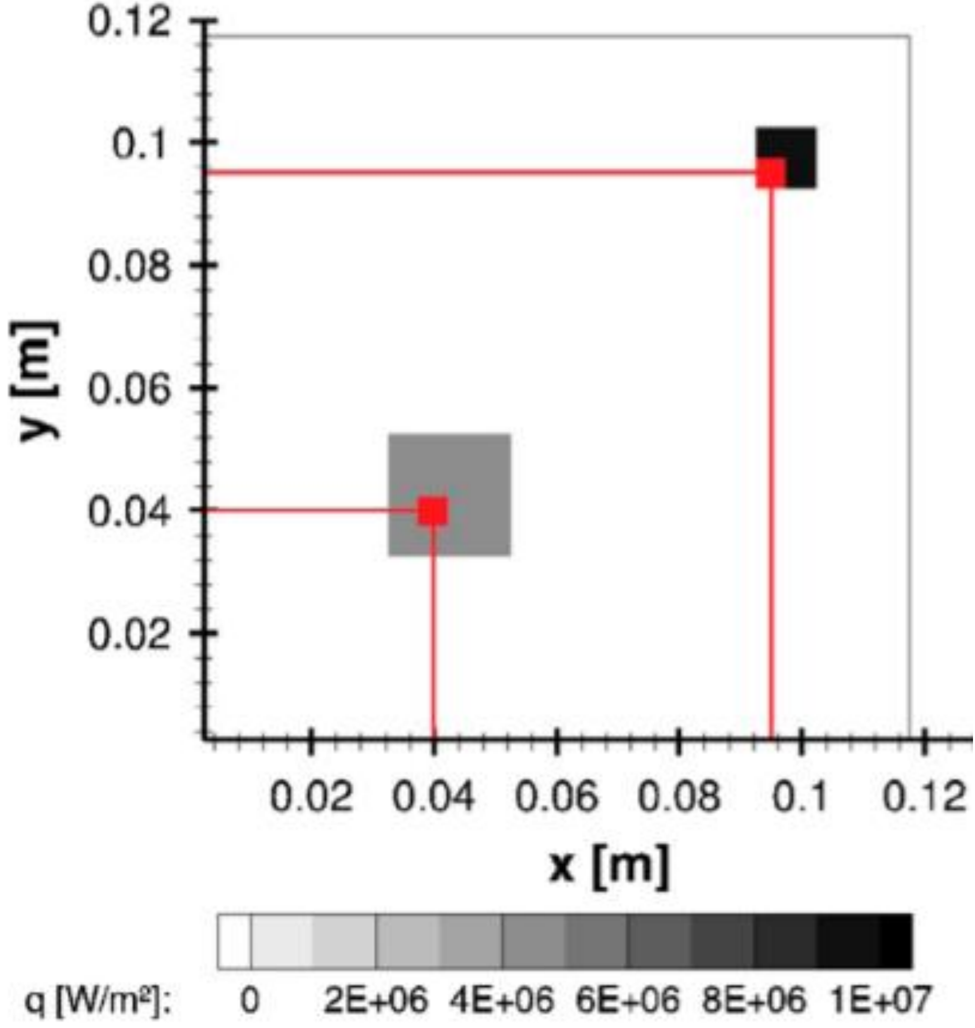
q_1	1E7 W/m^2
q_2	5E6 W/m^2
t_1	0.4 s
t_2	0.6 s

$x_{1,1}$	30 mm	$x_{2,1}$	90 mm
$x_{1,2}$	50 mm	$x_{2,2}$	100 mm
$y_{1,1}$	30 mm	$y_{2,1}$	90 mm
$y_{1,2}$	50 mm	$y_{2,2}$	100 mm

Example III – Inverse Problem Settings

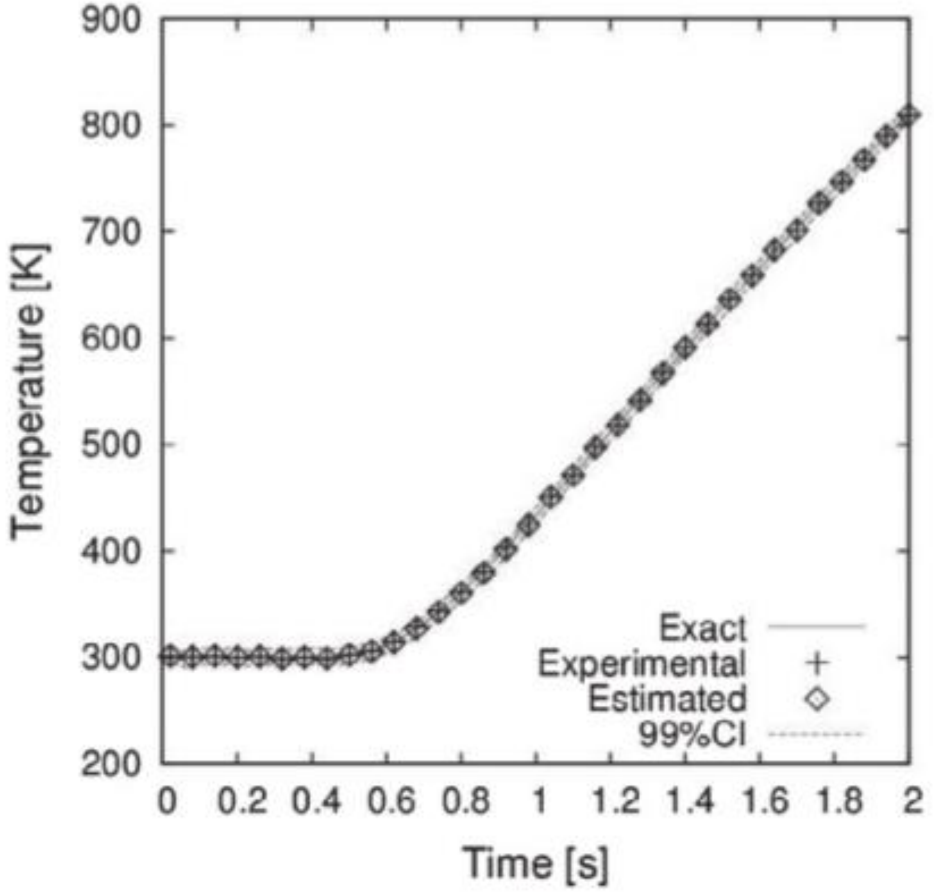
Exact (Reference) Heat Flux:

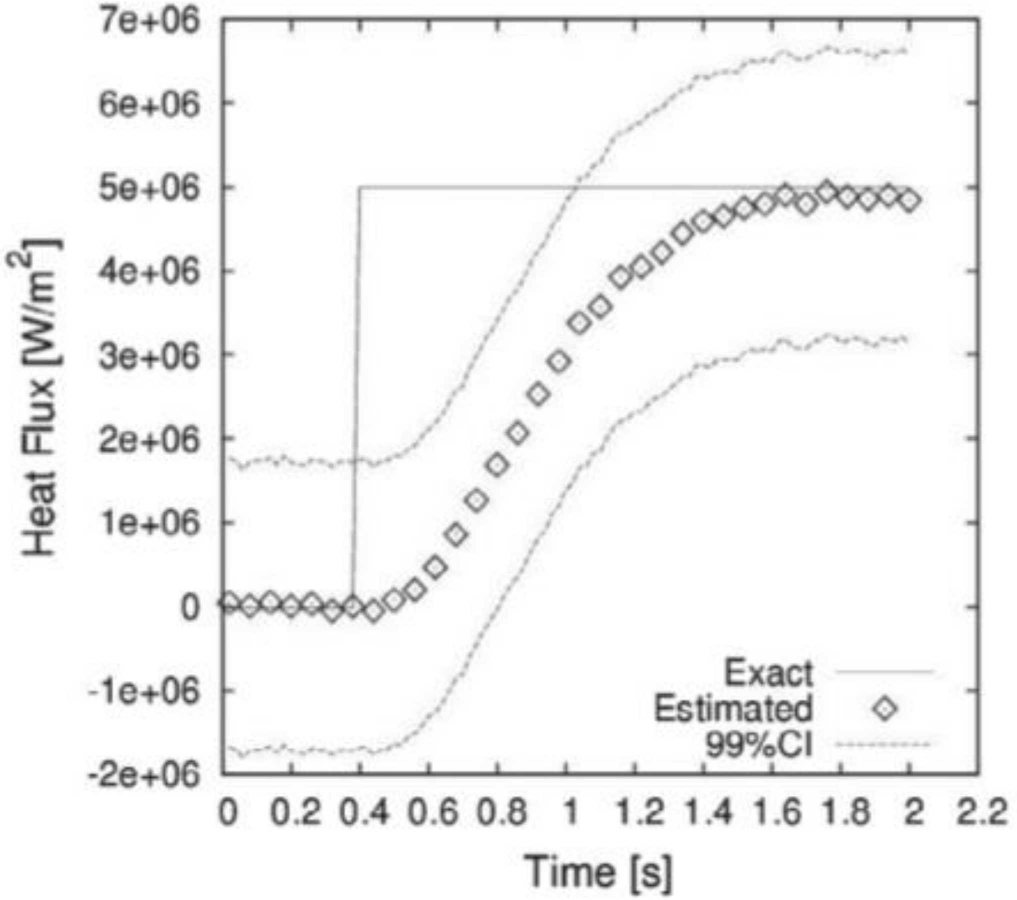
$$q(x,y,t)=\begin{cases} q_1, & \begin{array}{c} x_{1,1}\le x\le x_{1,2}, \\ y_{1,1}\le y\le y_{1,2}, \end{array} \quad t\ge t_0 \\ q_2 & \begin{array}{c} x_{2,1}\le x\le x_{2,2}, \\ y_{2,1}\le y\le y_{2,2}, \end{array} \quad t\ge t_0 \\ 0, & \text{caso contrário} \end{cases}$$



Example III – Results

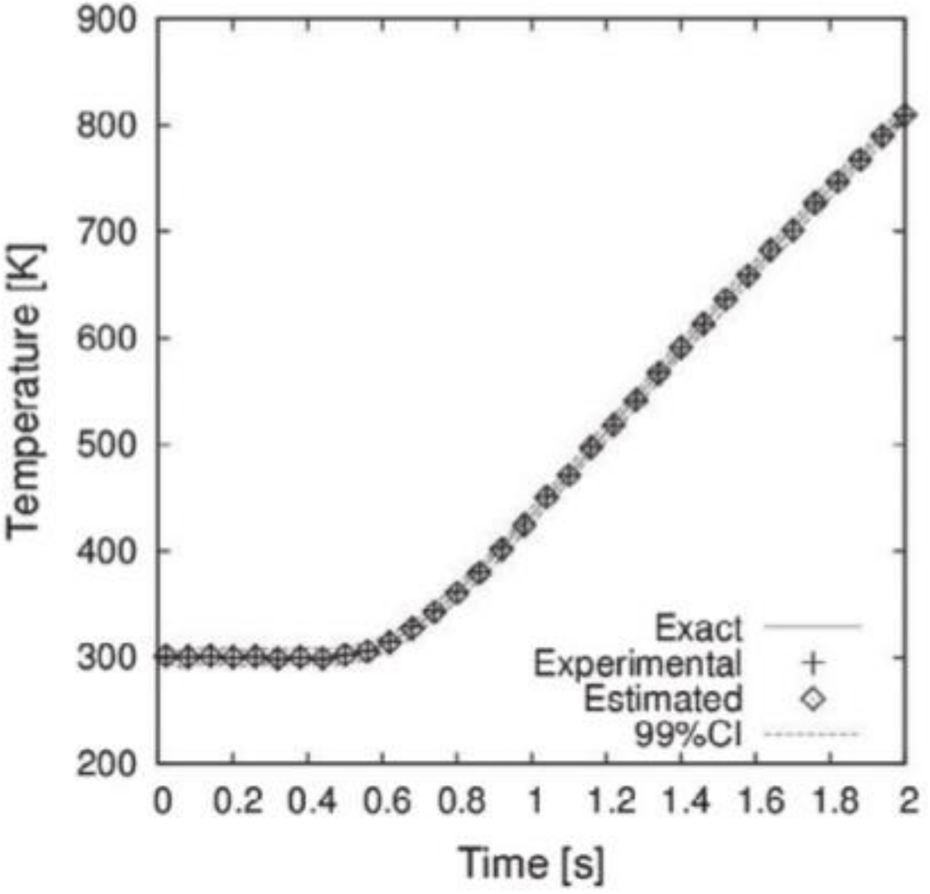
Hot Spot #1: Temperature and heat flux estimates

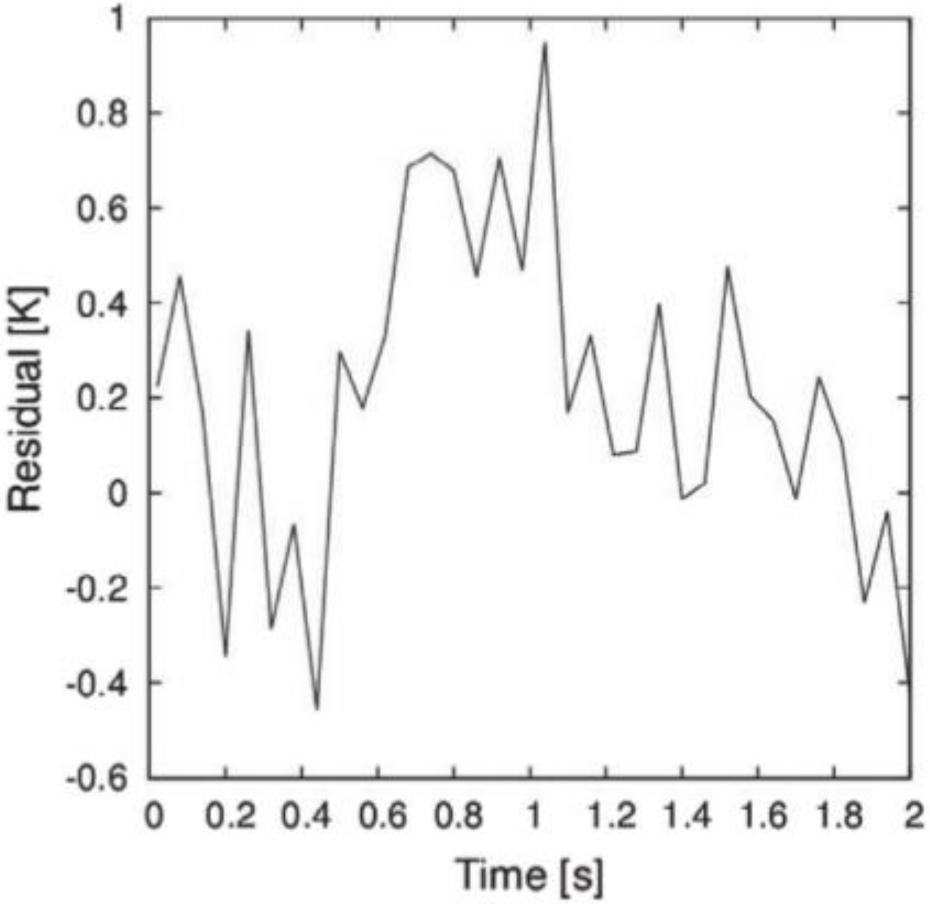




Example III – Results

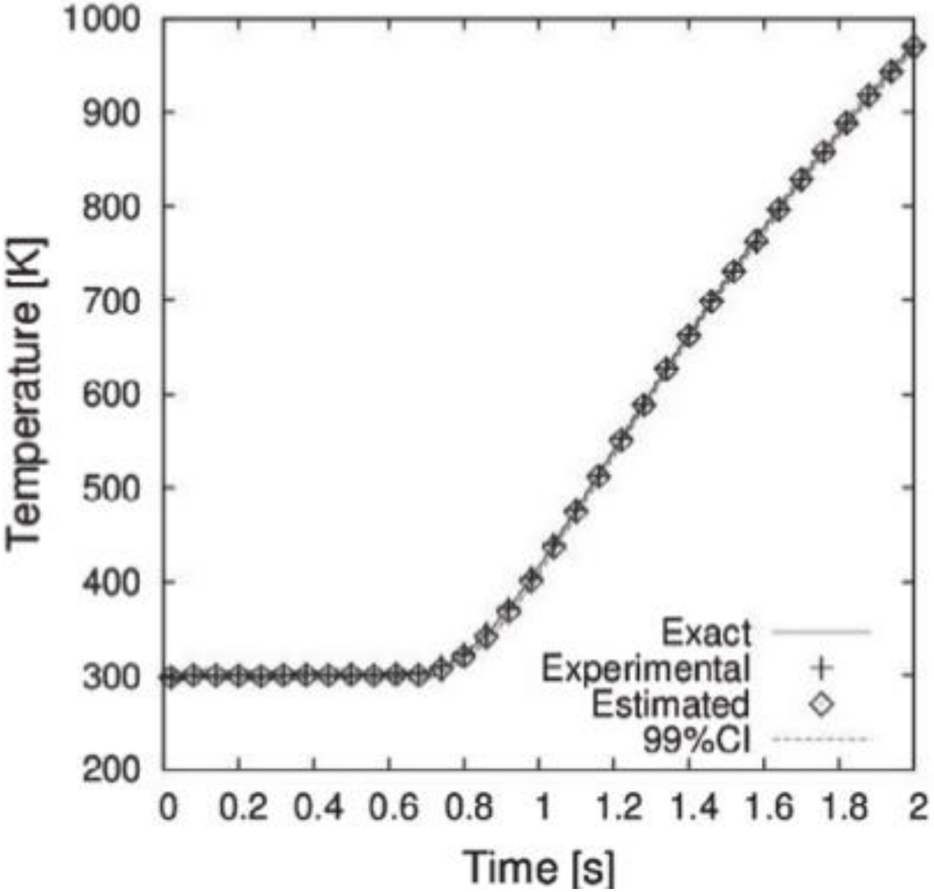
Hot Spot #1: Temperature estimates and residuals

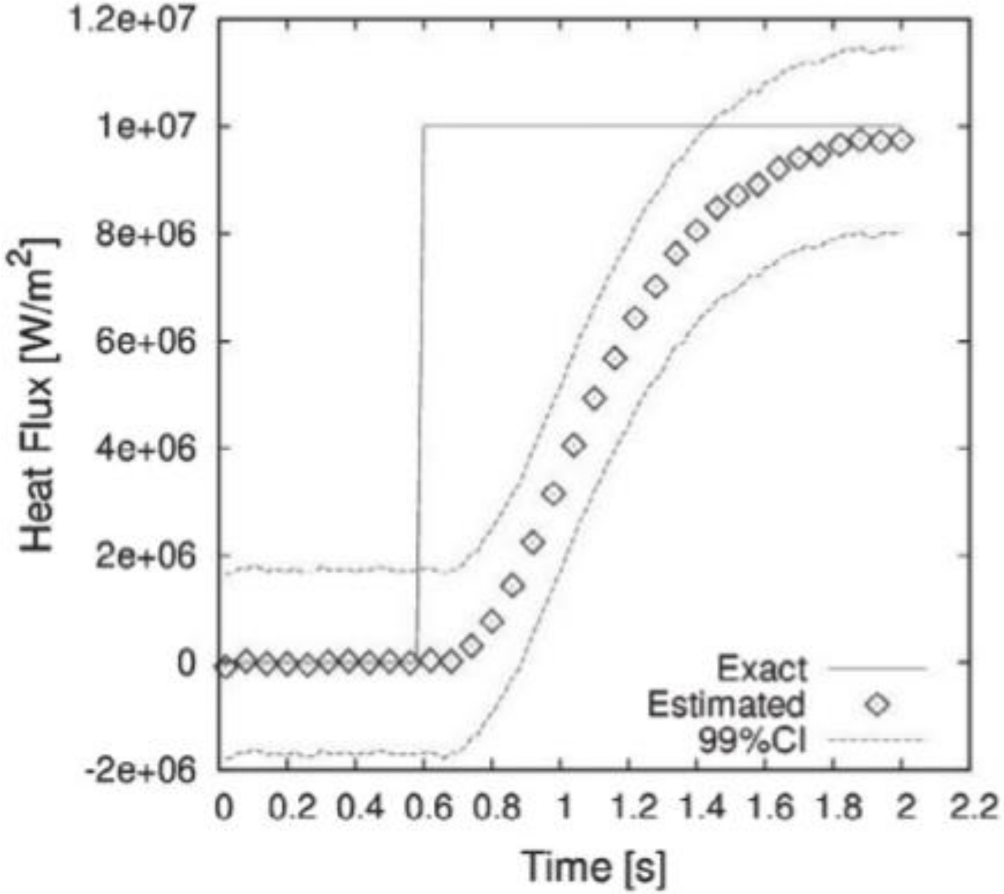




Example III – Results

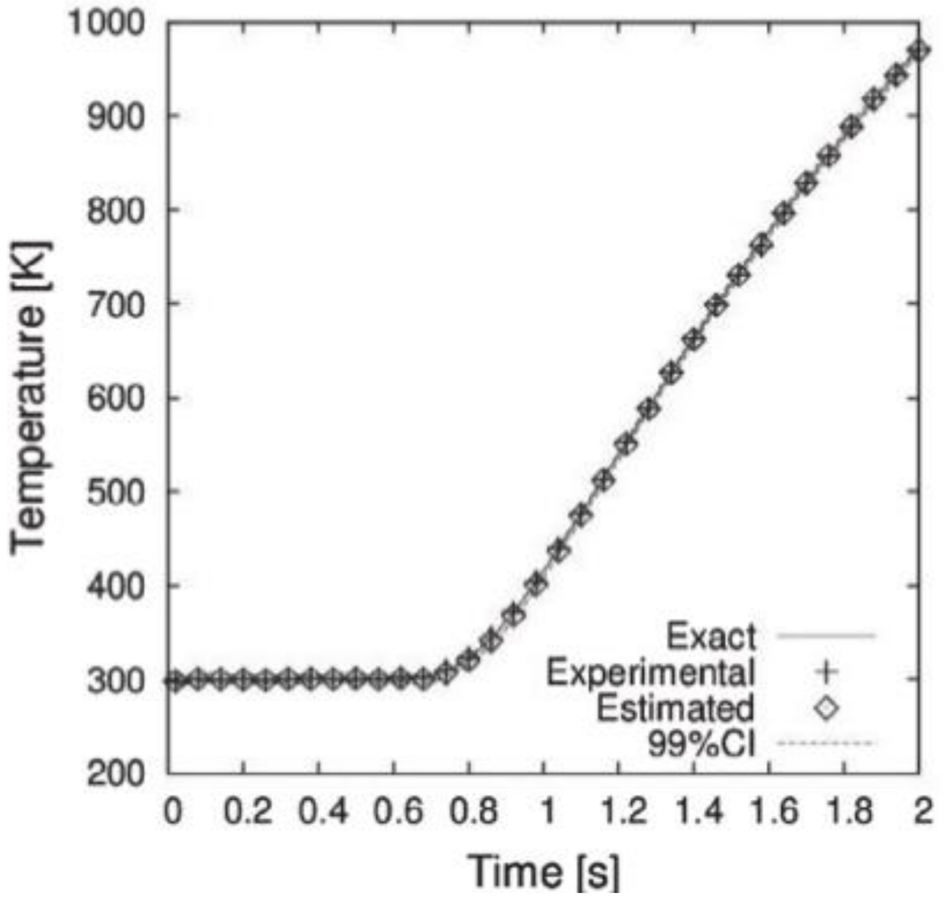
Hot Spot #2: Temperature and heat flux estimates

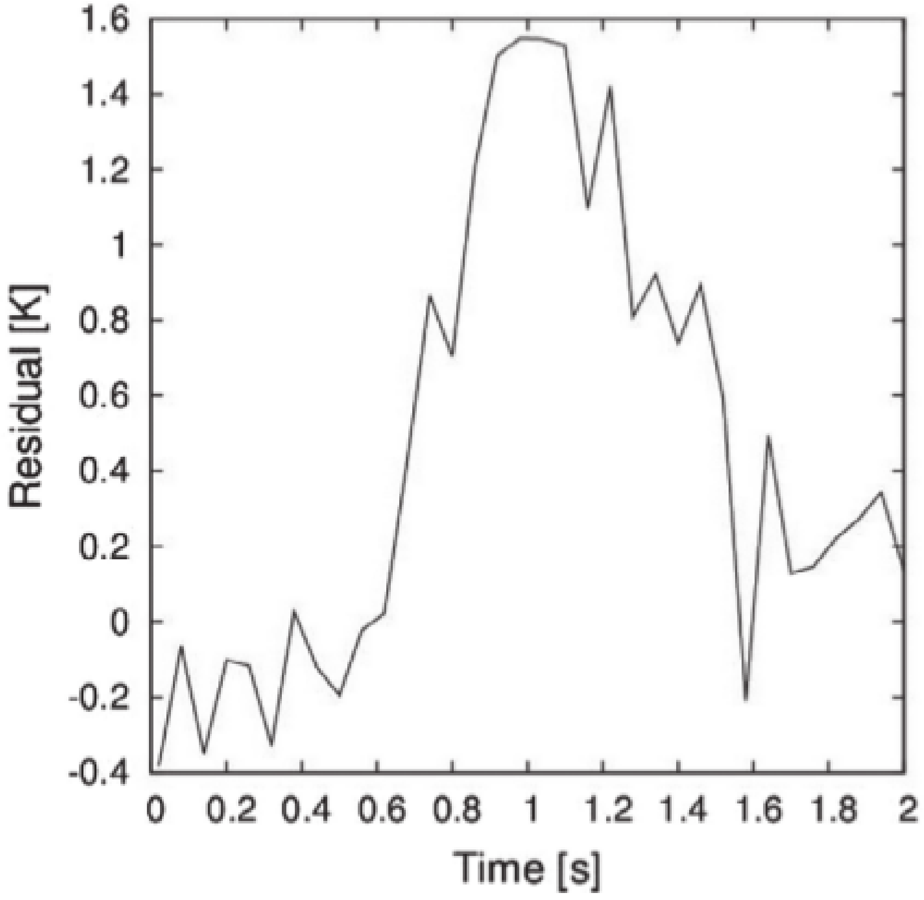




Example III – Results

Hot Spot #2: Temperature and heat flux estimates





Example III – Results

Computational time:

- Physical time: 2.0 s;
- Kalman filter: 300 s;
- **SSKF: 0.9 s;**

Only SSKF allows for real-time estimation;

- Majority of computational effort done at pre-processing;
- Recursive estimation: 2 matrix-vector multiplications.

The Kalman Filter – Extensions

- What if instead of

$$\boxed{\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{u}_{k-1} + \mathbf{w}_{k-1}} \qquad \boxed{\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{n}_k}$$

we had a nonlinear non Gaussian model:

$$\boxed{\mathbf{x}_k = \mathbf{f}_k\left(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1}\right)} \qquad \boxed{\mathbf{y}_k = \mathbf{h}\left(\mathbf{x}_k, \mathbf{n}_k\right)}$$

How does one perform sequential estimation with such models?

The Extended KF

- Idea: Linearize the model and apply classical KF equations:

$$\boxed{\mathbf{x}_k = \mathbf{f}_k\left(\hat{\mathbf{x}}_{k-1}^{+}, \mathbf{u}_{k-1}, \mathbf{w}_{n-1}\right) + \left.\frac{\partial \mathbf{f}_k}{\partial \mathbf{x}}\right|_{\hat{\mathbf{x}}_{k-1}^{+}} \left(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}^{+}\right) + \left.\frac{\partial \mathbf{f}_k}{\partial \mathbf{w}}\right|_{\hat{\mathbf{x}}_{k-1}^{+}} \mathbf{w}_{k-1}}$$

$$\boxed{\mathbf{y}_k = \mathbf{h}_k\left(\hat{\mathbf{x}}_{k}^{-}, \mathbf{n}_n\right) + \left.\frac{\partial \mathbf{h}_k}{\partial \mathbf{x}}\right|_{\hat{\mathbf{x}}_{k}^{-}} \left(\mathbf{x}_{k} - \hat{\mathbf{x}}_{k}^{-}\right) + \left.\frac{\partial \mathbf{h}_k}{\partial \mathbf{n}}\right|_{\hat{\mathbf{x}}_{k}^{-}} \mathbf{n}_{k}}$$

The Extended KF – Equations

Extended KF (EKF): Linearization + Kalman Filter:

- Linearization of both evolution and observation models.

Prediction:

- **Linearization:**

$$\boxed{\mathbf{F}_{k-1} = \left.\frac{\partial \mathbf{f}_k}{\partial \mathbf{x}}\right|_{\hat{\mathbf{x}}_{k-1}^{+}} \quad \text{and} \quad \mathbf{L}_{k-1} = \left.\frac{\partial \mathbf{f}_k}{\partial \mathbf{w}}\right|_{\hat{\mathbf{x}}_{k-1}^{+}}}$$

- **Prior Mean and Covariance**

$$\boxed{\hat{\mathbf{x}}_n^{-} = \mathbf{f}_k\left(\hat{\mathbf{x}}_{k-1}^{+}, \mathbf{u}_{k-1}, \mathbf{0}\right) \quad \text{and} \quad \mathbf{P}_n^{-} = \mathbf{F}_{k-1}\mathbf{P}_{k-1}^{+}\mathbf{F}_{k-1}^{T} + \mathbf{L}_{k-1}\mathbf{Q}_{k-1}\mathbf{L}_{k-1}^{T}}$$

The Extended KF – Equations

Update:

- **Linearization**

$$\boxed{\mathbf{H}_k = \left.\frac{\partial \mathbf{h}_k}{\partial \mathbf{x}}\right|_{\hat{\mathbf{x}}_k^-} \quad \text{and} \quad \mathbf{M}_k = \left.\frac{\partial \mathbf{h}_k}{\partial \mathbf{n}}\right|_{\hat{\mathbf{x}}_k^-}}$$

- **Calculation of Intermediate Covariance Matrices**

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{M}_k \mathbf{R}_k \mathbf{M}_k^T \right)^{-1}$$

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \left[\mathbf{y}_k - \mathbf{h}_k \left(\hat{\mathbf{x}}_k^-, \mathbf{0} \right) \right] \quad \text{and} \quad \mathbf{P}_k^+ = \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k \right) \mathbf{P}_k^-$$

Example IV

Free falling of a particle:

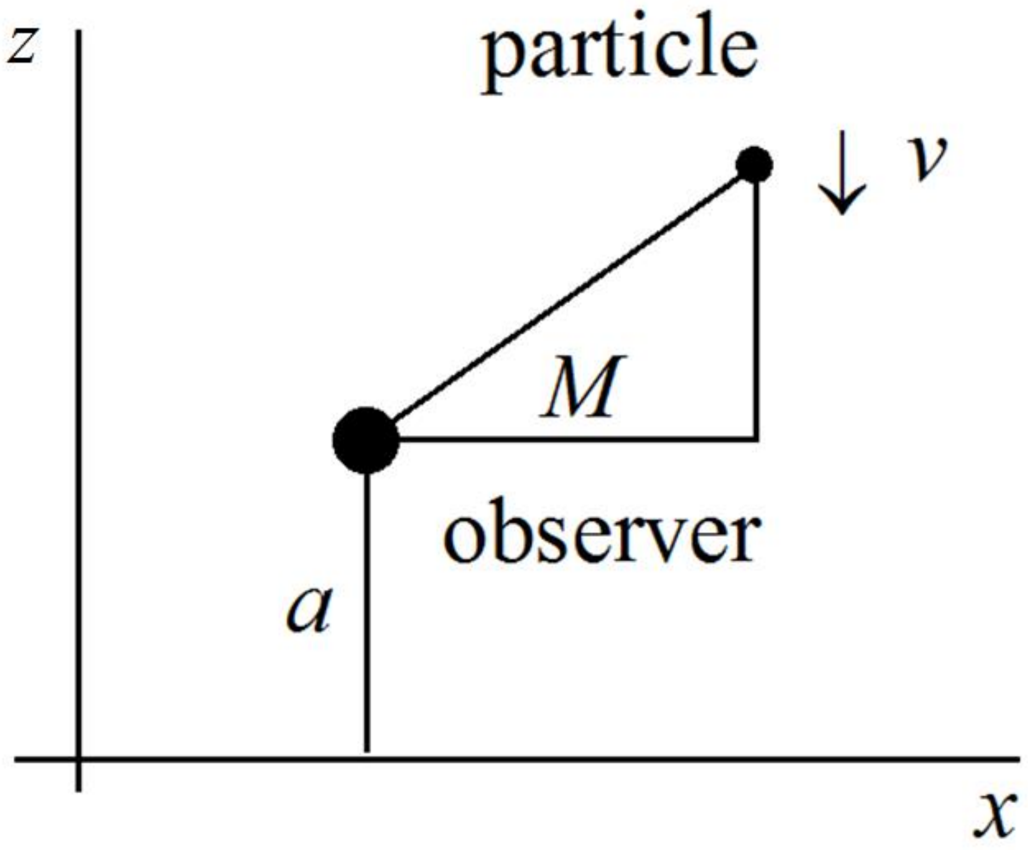
State variables:

- Altitude (x_1);
- Velocity (x_2);
- Ballistic Coefficient (x_3);

Observation variable:

- Observed distance:

$$\sqrt{M^2 + \left[x_1(t) - a\right]^2}$$



Example IV

Free falling of a particle:

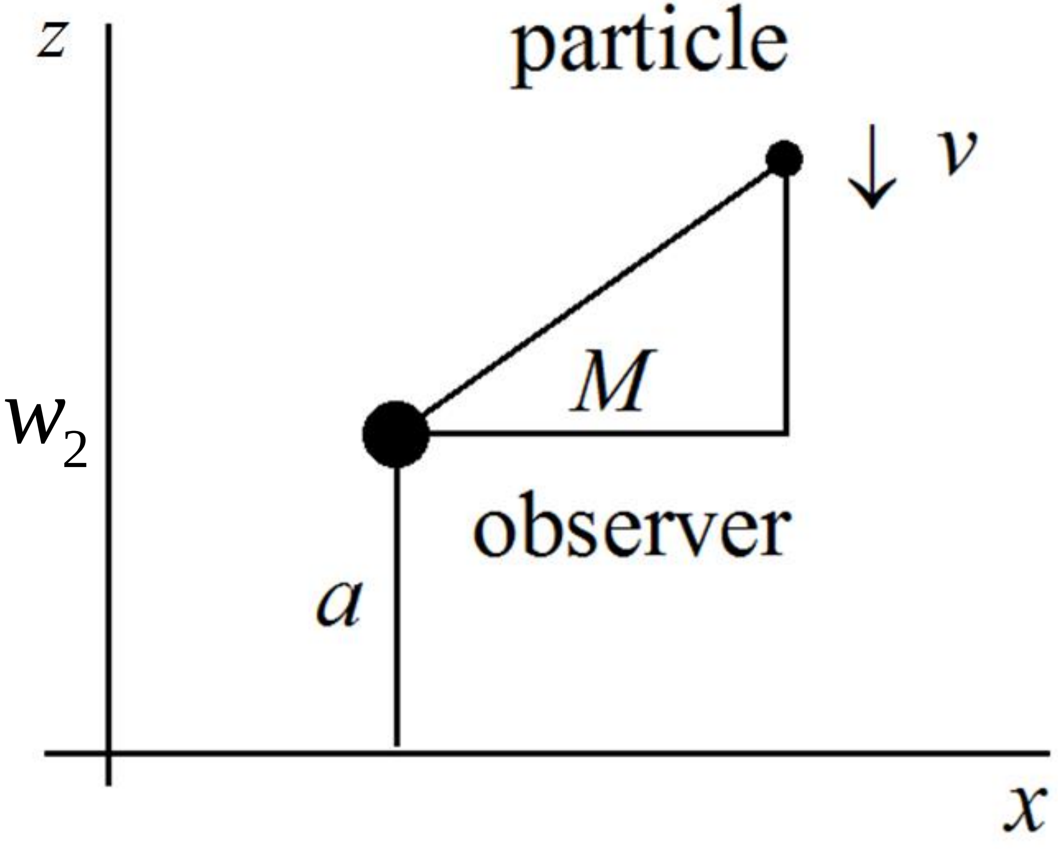
Evolution-Observation Model

$$\dot{x}_1 = x_2 + w_1$$

$$\dot{x}_2 = {}^1\!/_2\,\rho_0 \exp\left[-{}^{x_1}\!/_k\right] x_2^2 x_3 - g + w_2$$

$$\dot{x}_3 = w_3$$

$$y(t_k) = \sqrt{M^2 + \left[x_1(t_k) - a\right]^2} + v_k$$



Example IV

Free falling of a particle:

Input data:

$$\rho_0 = 2 \text{ lbs}^2/\text{ft}^4;$$

$$g \quad = 32.2 \text{ ft/s}^2;$$

$$k \quad = 2\times 10^4 \text{ ft};$$

$$M = 10^5 \text{ ft};$$

$$\text{a} \quad = 10^5 \text{ ft}.$$

Noise model:

$$\mathrm{E}\left[w_i^2(t)\right] = 0, \quad i = 1, 2, 3$$

$$\mathrm{E}\left[v_k^2\right] = 10^4 \text{ ft}^2$$

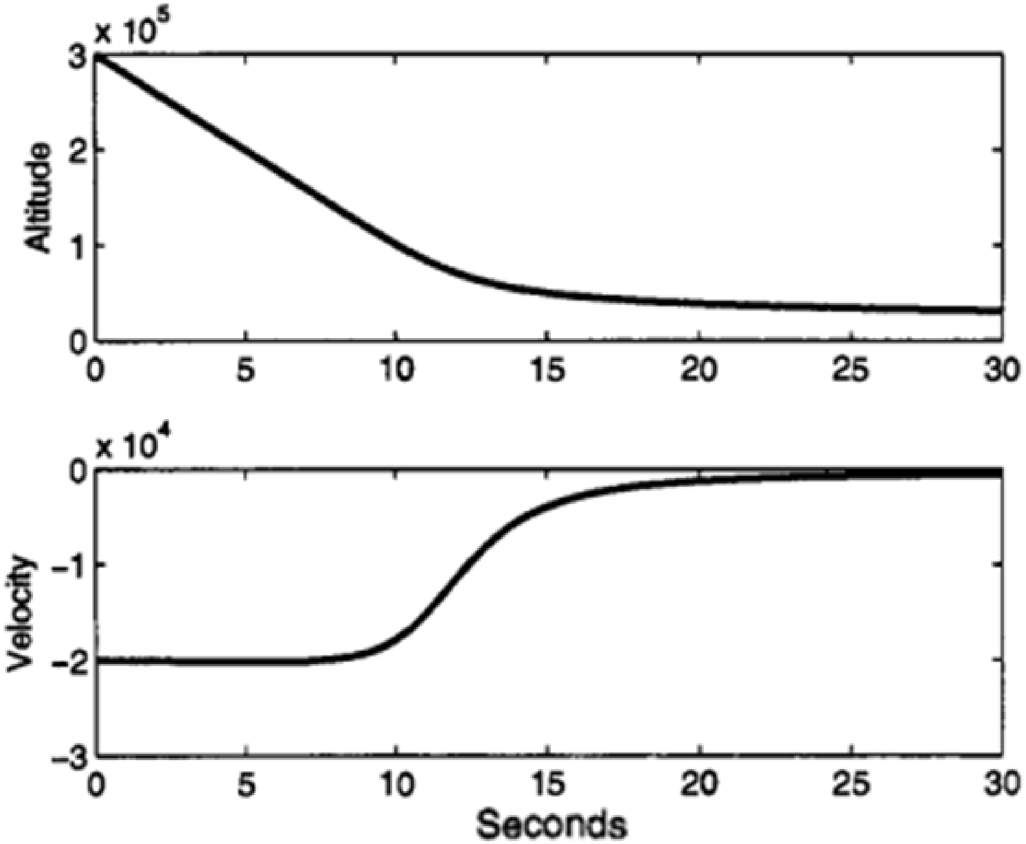
Initial State:

$$\hat{\mathbf{x}}_0^+ = \mathbf{x}_0 = \begin{bmatrix} 3\times 10^5 \\ -2\times 10^4 \\ 10^{-3} \end{bmatrix} \quad \mathbf{P}_0^+ = \mathrm{diag}\begin{bmatrix} 10^6 \\ 4\times 10^6 \\ 10 \end{bmatrix}$$

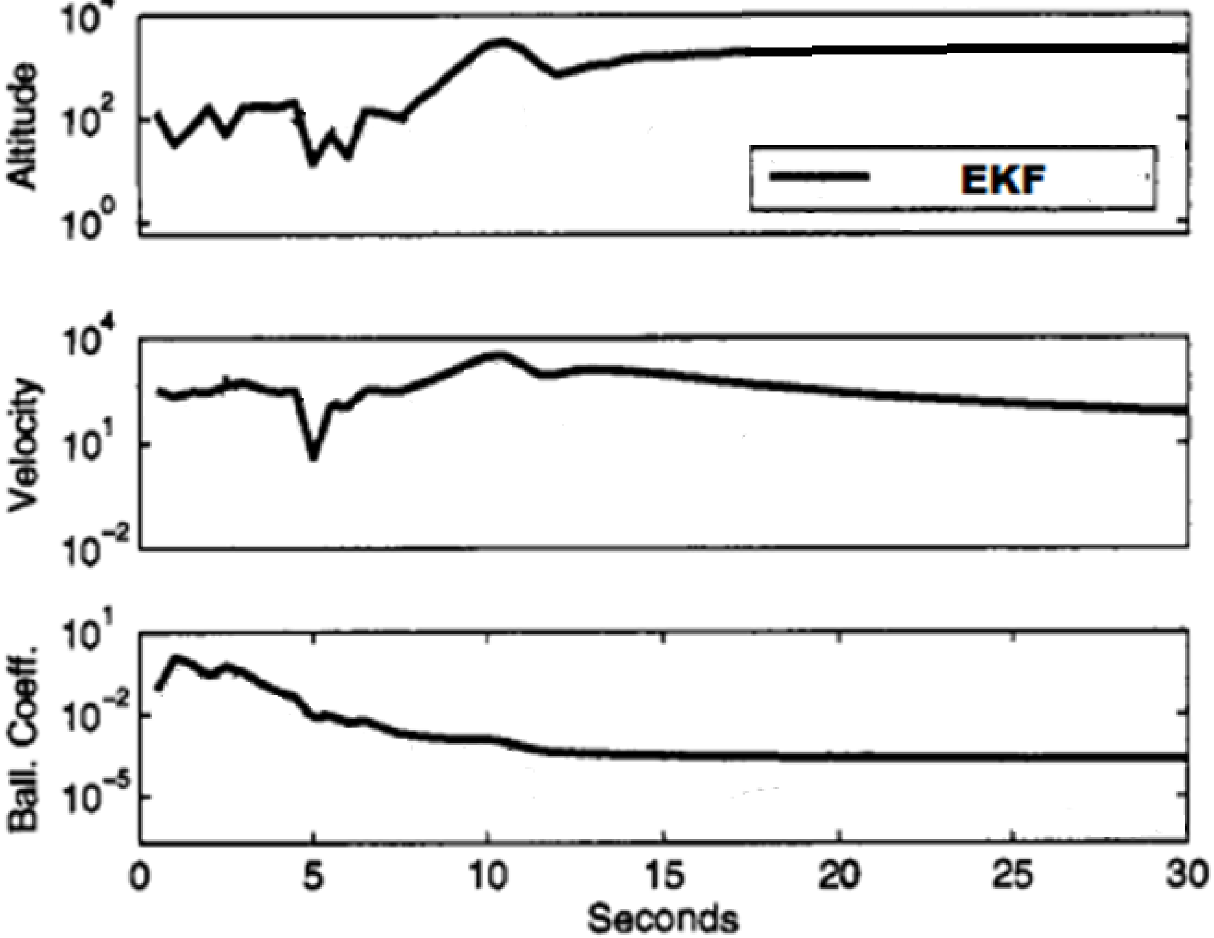
Example IV

Free falling of a particle:

Reference values:



Estimation error:



Higher order methods

EKF properties:

- Sequential nonlinear estimation algorithm of choice;
- Non-intrusive;
- Up for parallelization;
- Hard to accelerate theoretically (there is no “SSEKF”).

What if nonlinearities are not sufficiently captured?

- 1st order approximations might be insufficient;
- EKF might lead to unreliable estimates.

Higher order methods

Example: Polar to Rectangular mapping:

$$\boxed{\mathbf{x} = \begin{bmatrix} r \\ \theta \end{bmatrix} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix}}$$

Goal: Find **y** statistics, given

- Nonlinear mapping;
- **x** is uncorrelated;
- Symmetric pdfs.

$$\boxed{\bar{\mathbf{x}} = \begin{bmatrix} \bar{r} \\ \bar{\theta} \end{bmatrix} = \begin{bmatrix} 1 \\ \pi/2 \end{bmatrix} \quad \text{and} \quad \mathbf{y} = \mathbf{h}(\mathbf{x})}$$

Higher order methods

- 1st order expansion around mean of **x**:

$$\bar{\mathbf{y}} = \mathrm{E}\left[\mathbf{h}\left(\mathbf{x}\right)\right] \simeq \mathrm{E}\left[\mathbf{h}\left(\bar{\mathbf{x}}\right) + \left.\frac{\partial \mathbf{h}}{\partial \mathbf{x}}\right|_{\bar{\mathbf{x}}}\left(\mathbf{x} - \bar{\mathbf{x}}\right)\right]$$

- Thus

$$\boxed{\bar{\mathbf{y}} \simeq \mathbf{h}\left(\bar{\mathbf{x}}\right) + \left.\frac{\partial \mathbf{h}}{\partial \mathbf{x}}\right|_{\bar{\mathbf{x}}} \mathrm{E}\left[\mathbf{x} - \bar{\mathbf{x}}\right] \quad \therefore \quad \bar{\mathbf{y}} \simeq \mathbf{h}\left(\bar{\mathbf{x}}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}}$$

Higher order methods

1st component:

$$\begin{aligned}\bar{y}_1 &= \mathrm{E}\left[r\cos\theta\right]\\&= \mathrm{E}\left[\left(\bar{r}+\tilde{r}\right)\cos\left(\bar{\theta}+\tilde{\theta}\right)\right]\\&= \mathrm{E}\left[\left(\bar{r}+\tilde{r}\right)\left(\cos\bar{\theta}\cos\tilde{\theta}-\sin\bar{\theta}\sin\tilde{\theta}\right)\right]\\&= \bar{r}\cos\bar{\theta}\\&= 0\end{aligned}$$

Assuming:

$$\mathbf{x} = \begin{bmatrix} r \\ \theta \end{bmatrix} = \begin{bmatrix} \bar{r}+\tilde{r} \\ \bar{\theta}+\tilde{\theta} \end{bmatrix}$$

Thus: $\bar{y}_1 = \bar{r}\cos\bar{\theta} \quad \therefore \quad \bar{y}_1 = 0$

Higher order methods

2nd component:

$$\bar{y}_2 = \mathrm{E}\left[r \sin\theta\right]$$
$$= \mathrm{E}\left[(\bar{r}+\tilde{r})\sin\left(\bar{\theta}+\tilde{\theta}\right)\right]$$
$$= \mathrm{E}\left[(\bar{r}+\tilde{r})\left(\sin\bar{\theta}\cos\tilde{\theta} - \cos\bar{\theta}\sin\tilde{\theta}\right)\right]$$
$$= \bar{r}\sin\bar{\theta}\,\mathrm{E}\left[\cos\tilde{\theta}\right]$$
$$= \mathrm{E}\left[\cos\tilde{\theta}\right]$$

Assuming:

$$\mathbf{x} = \begin{bmatrix} r \\ \theta \end{bmatrix} = \begin{bmatrix} \bar{r}+\tilde{r} \\ \bar{\theta}+\tilde{\theta} \end{bmatrix}$$

Thus: $\bar{y}_2 = \mathrm{E}\left[\cos\tilde{\theta}\right]$ but $\bar{y}_2 = 1$???

Higher order methods

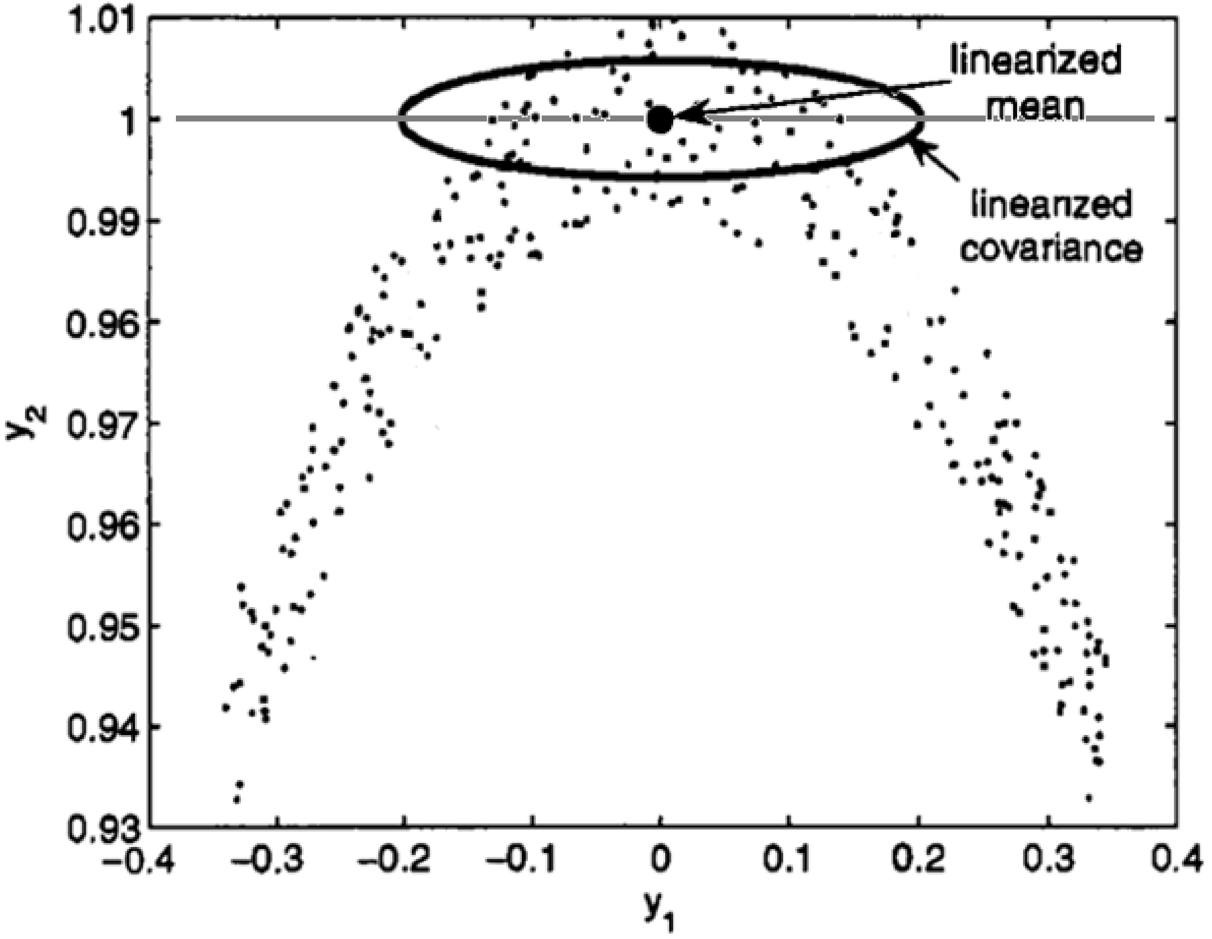
If we assume uniform distribution:

$$\boxed{\theta \sim U\left[-\theta_m, \theta_m\right]}$$

It follows that

$$\boxed{\mathrm{E}\left[\cos\tilde{\theta}\right] = \frac{\sin\theta_m}{\theta_m} < 1}$$

The first order approximation is incorrect!!



Higher order methods

How to improve this estimate?

In other words, we seek better ways to propagate

$$\bar{\mathbf{x}} = \mathrm{E}\left[\mathbf{x}\right] \text{ and } \mathbf{P} = \mathrm{cov}\left[\mathbf{x}\right]$$

through a nonlinear mapping

$$\mathbf{y} = \mathbf{h}\left(\mathbf{x}\right)$$

in order to obtain

$$\bar{\mathbf{y}} = \mathrm{E}\left[\mathbf{y}\right] \text{ and } \mathbf{P}_y = \mathrm{cov}\left[\mathbf{y}\right]$$

Higher order methods

Solution: Unscented transform.

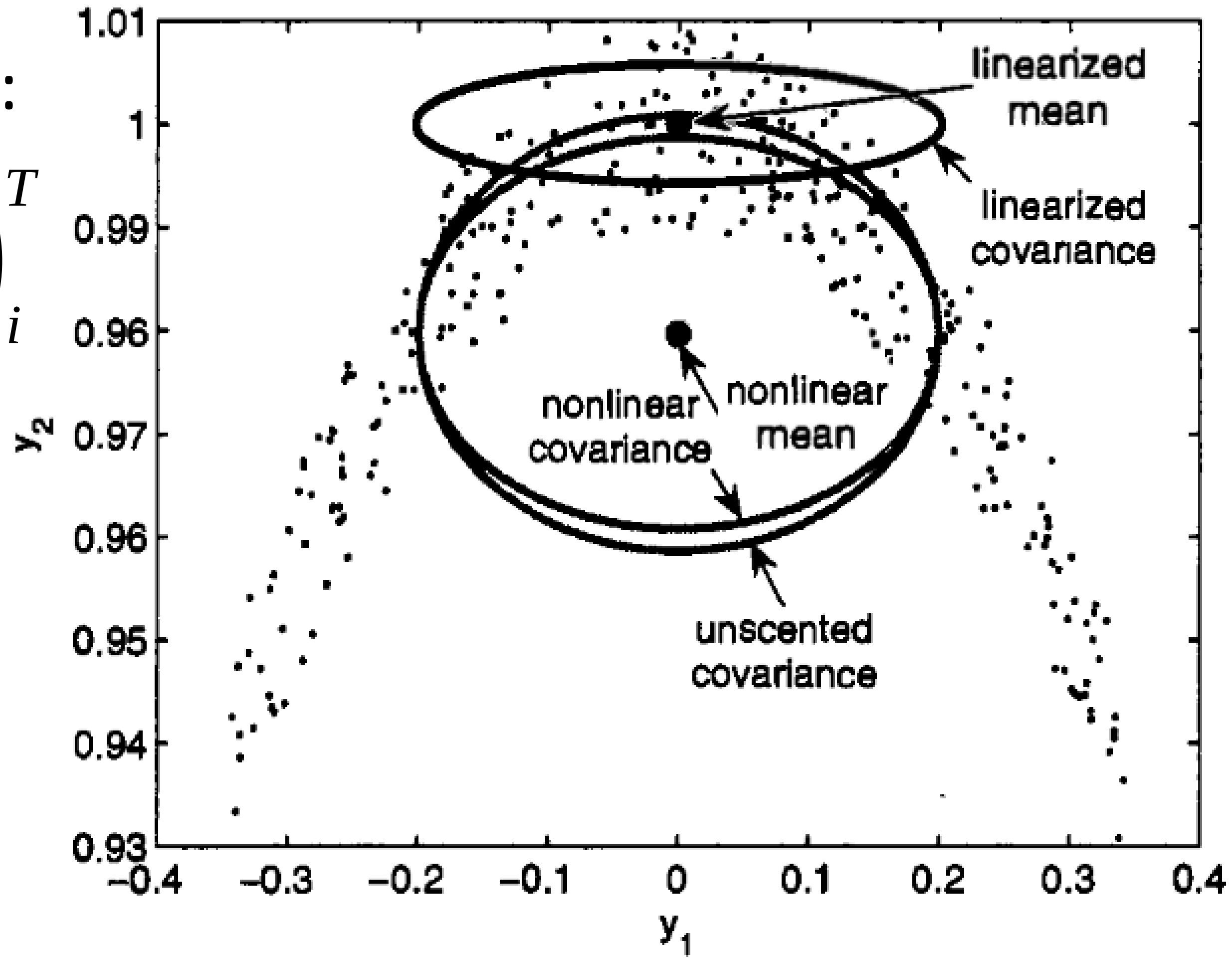
1st step (2n samples or sigma points):

$$\mathbf{x}^{(i)} = \bar{\mathbf{x}} + \tilde{\mathbf{x}}^{(i)}, \text{with } \tilde{\mathbf{x}}^{(i)} = \pm\left(\sqrt{n\mathbf{P}}\right)_i^T$$

- n is the size of **x** vector;
- i subscript: ith row of matrix.

2nd step (Mapping of samples):

$$\mathbf{y}^{(i)} = \mathbf{h}\left(\mathbf{x}^{(i)}\right)$$

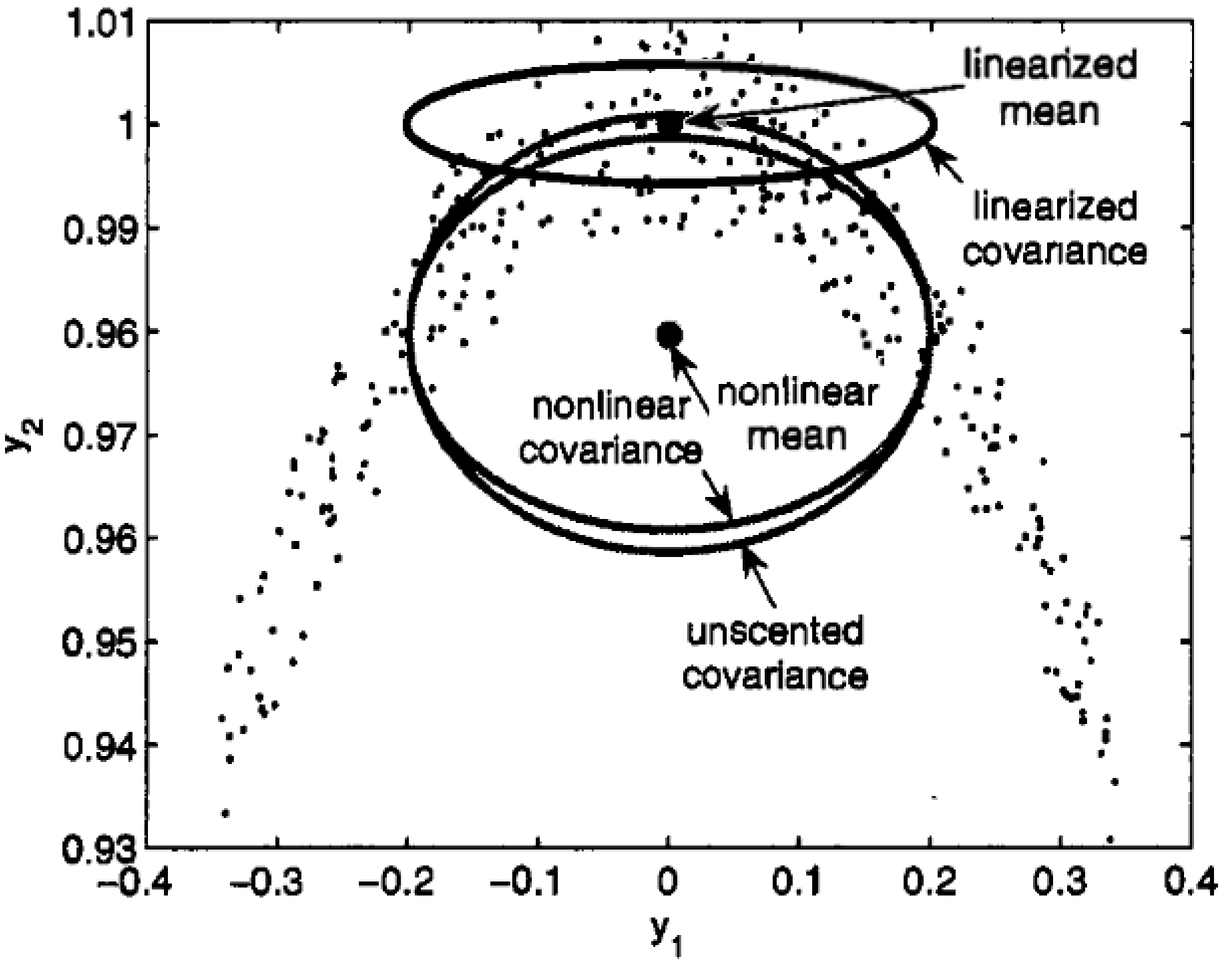


Higher order methods

Solution: Unscented transform.

3rd step (Averaging of mappings):

$$\boxed{\begin{gathered}\bar{\mathbf{y}} = \tfrac{1}{2n}\sum_{i=1}^{2n}\mathbf{y}^{(i)} \\ \text{and} \\ \mathbf{P}_y = \tfrac{1}{2n}\sum_{i=1}^{2n}\Delta\mathbf{y}\Delta\mathbf{y}^T\end{gathered}}$$



OBS: $\Delta\mathbf{y}^{(i)} = \mathbf{y}^{(i)} - \bar{\mathbf{y}}$

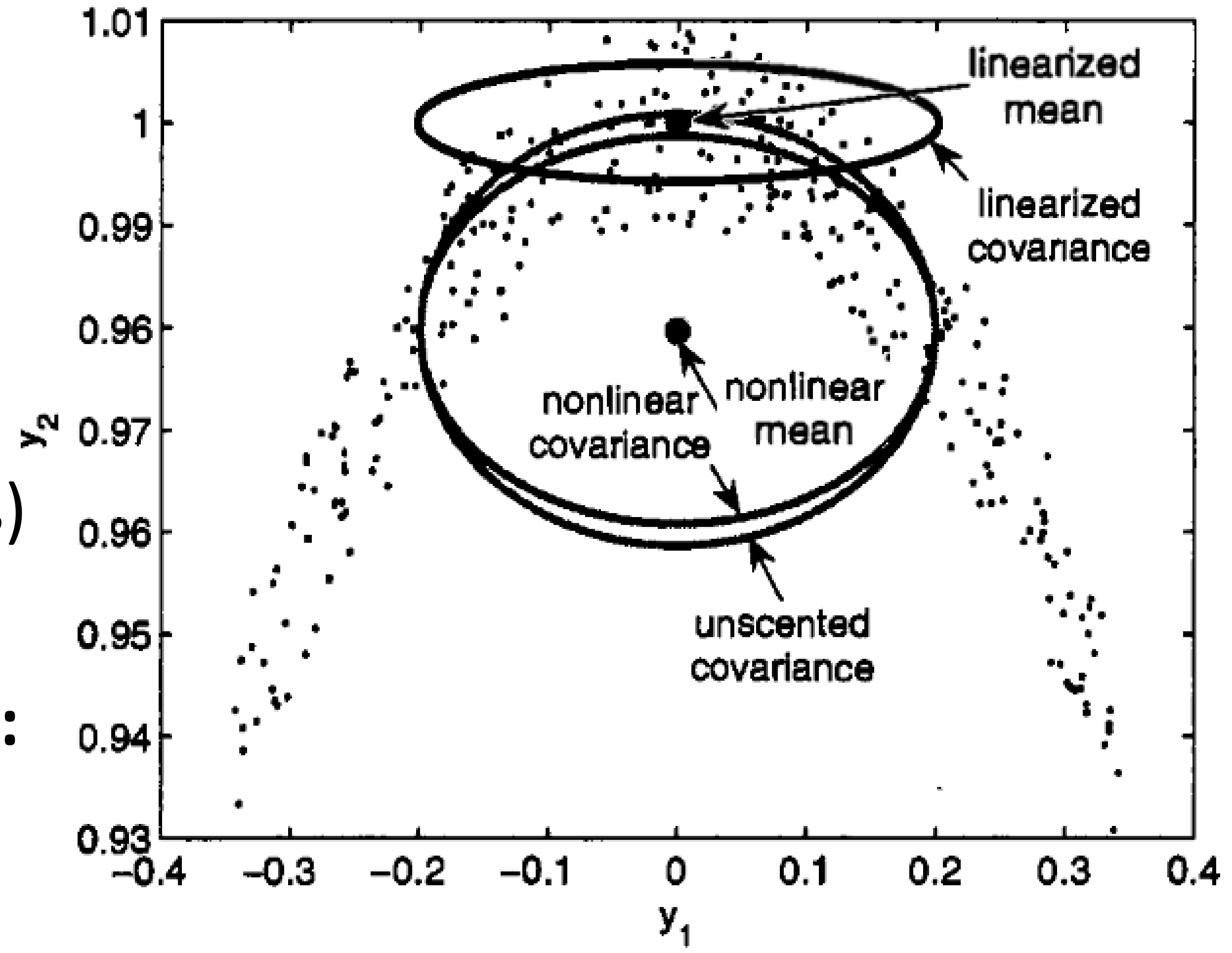
Higher order methods

Unscented transform:

- 3rd order accurate;
- Nonintrusive;
- No extra matrices required;
- Sq. root of **P**: Cholesky decomp.;
- Deterministic sampling (2n samples)

So, for heavily nonlinear problems:

UT-based Kalman filter



The Unscented KF – Equations

Unscented KF (UKF): Unscented transform (UT) + Kalman Filter:

- 1 UT at prediction stage + 1 UT at update stage.

Prediction:

- **Sampling of Sigma Points**

$$\hat{\mathbf{x}}_{n-1}^{(i)} = \hat{\mathbf{x}}_{n-1}^{+} \pm \left(\sqrt{n\mathbf{P}_{k-1}^{+}} \right)_{i}^{T} \quad \text{and} \quad \hat{\mathbf{x}}_{n}^{(i)} = \mathbf{f}\left(\hat{\mathbf{x}}_{n-1}^{(i)}, \mathbf{u}_{n} \right)$$

- **Prior Mean and Covariance**

$$\boxed{\hat{\mathbf{x}}_{n}^{-} = \frac{1}{2n} \sum_{i=1}^{2n} \hat{\mathbf{x}}_{n-1}^{(i)} \quad \text{and} \quad \mathbf{P}_{n}^{-} = \frac{1}{2n} \sum_{i=1}^{2n} \Delta\hat{\mathbf{x}}_{n}^{(i)} \Delta\hat{\mathbf{x}}_{n}^{(i),T} + \mathbf{Q}_{n}}$$

The Unscented KF – Equations

Update:

- **Sampling of sigma points**

$$\hat{\mathbf{x}}_n^{(i)} = \hat{\mathbf{x}}_n^- \pm \left(\sqrt{n\mathbf{P}_k^-}\right)_i^T \quad \text{and} \quad \hat{\mathbf{y}}_n^{(i)} = \mathbf{h}\left(\hat{\mathbf{x}}_n^{(i)}\right)$$

- **Calculation of Intermediate Covariance Matrices**

$$\hat{\mathbf{y}}_n = \frac{1}{2n}\sum_{i=1}^{2n}\hat{\mathbf{y}}_n^{(i)}$$

$$\mathbf{P}_y = \frac{1}{2n}\sum_{i=1}^{2n}\Delta\hat{\mathbf{y}}_n^{(i)}\Delta\hat{\mathbf{y}}_n^{(i),T} + \mathbf{R}_n \quad \text{and} \quad \mathbf{P}_{xy} = \frac{1}{2n}\sum_{i=1}^{2n}\Delta\hat{\mathbf{x}}_n^{(i)}\Delta\hat{\mathbf{y}}_n^{(i),T}$$

The Unscented KF – Equations

Update:

- **Posterior Mean and Covariance:**

$$\mathbf{K}_n = \mathbf{P}_{xy}\mathbf{P}_y^{-1}$$

$$\boxed{\hat{\mathbf{x}}_n^+ = \hat{\mathbf{x}}_n^- + \mathbf{K}_n\left[\mathbf{y}_n - \hat{\mathbf{y}}_n\right] \quad \text{and} \quad \mathbf{P}_n^+ = \mathbf{P}_n^- - \mathbf{K}_n\mathbf{P}_y\mathbf{K}_n^T}$$

Example V

Free falling of a particle – revisited:

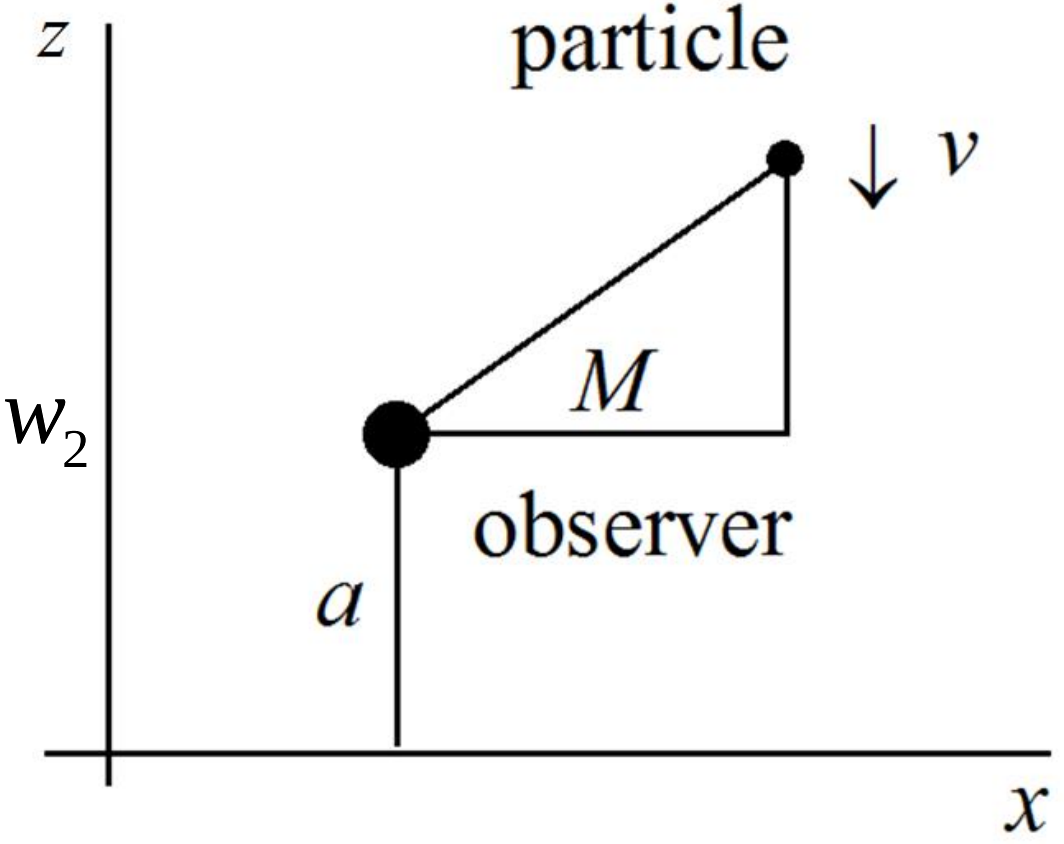
Evolution-Observation Model

$$\dot{x}_1 = x_2 + w_1$$

$$\dot{x}_2 = \tfrac{1}{2}\rho_0 \exp\left[-\tfrac{x_1}{k}\right] x_2^2 x_3 - g + w_2$$

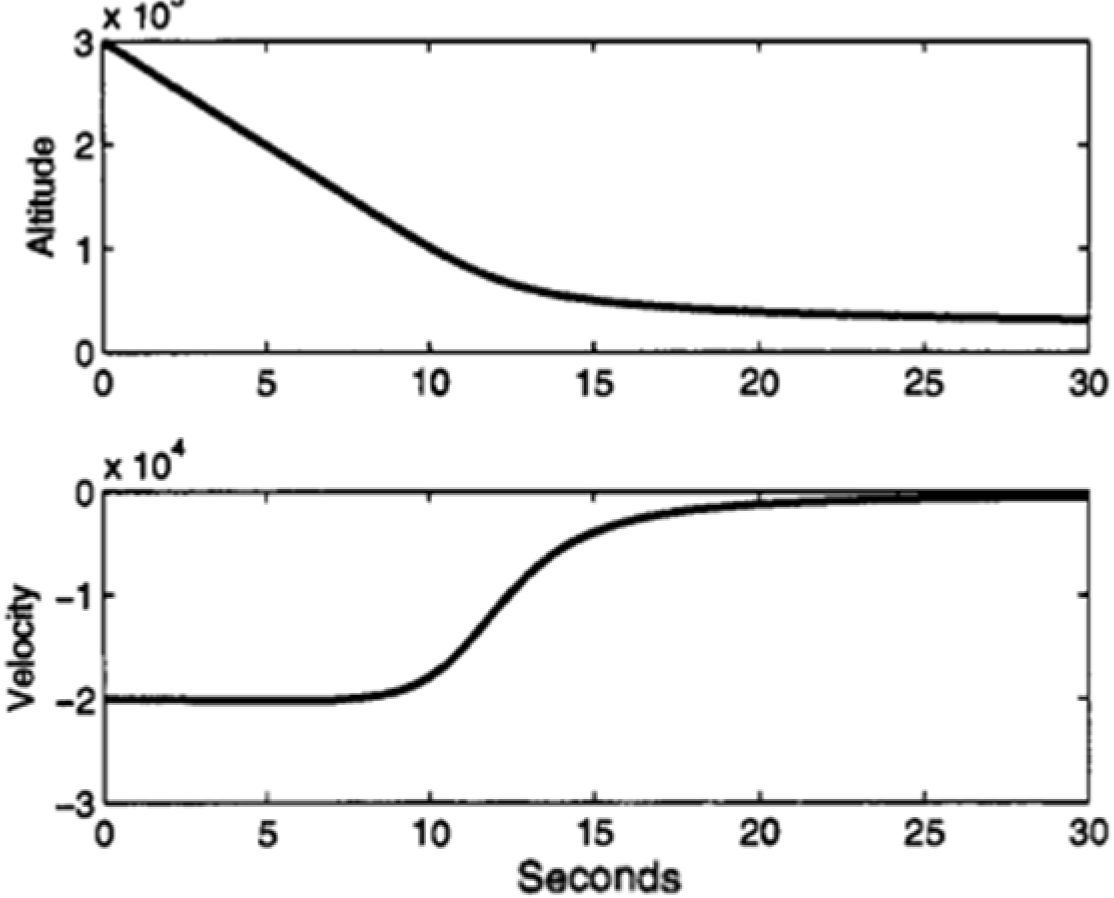
$$\dot{x}_3 = w_3$$

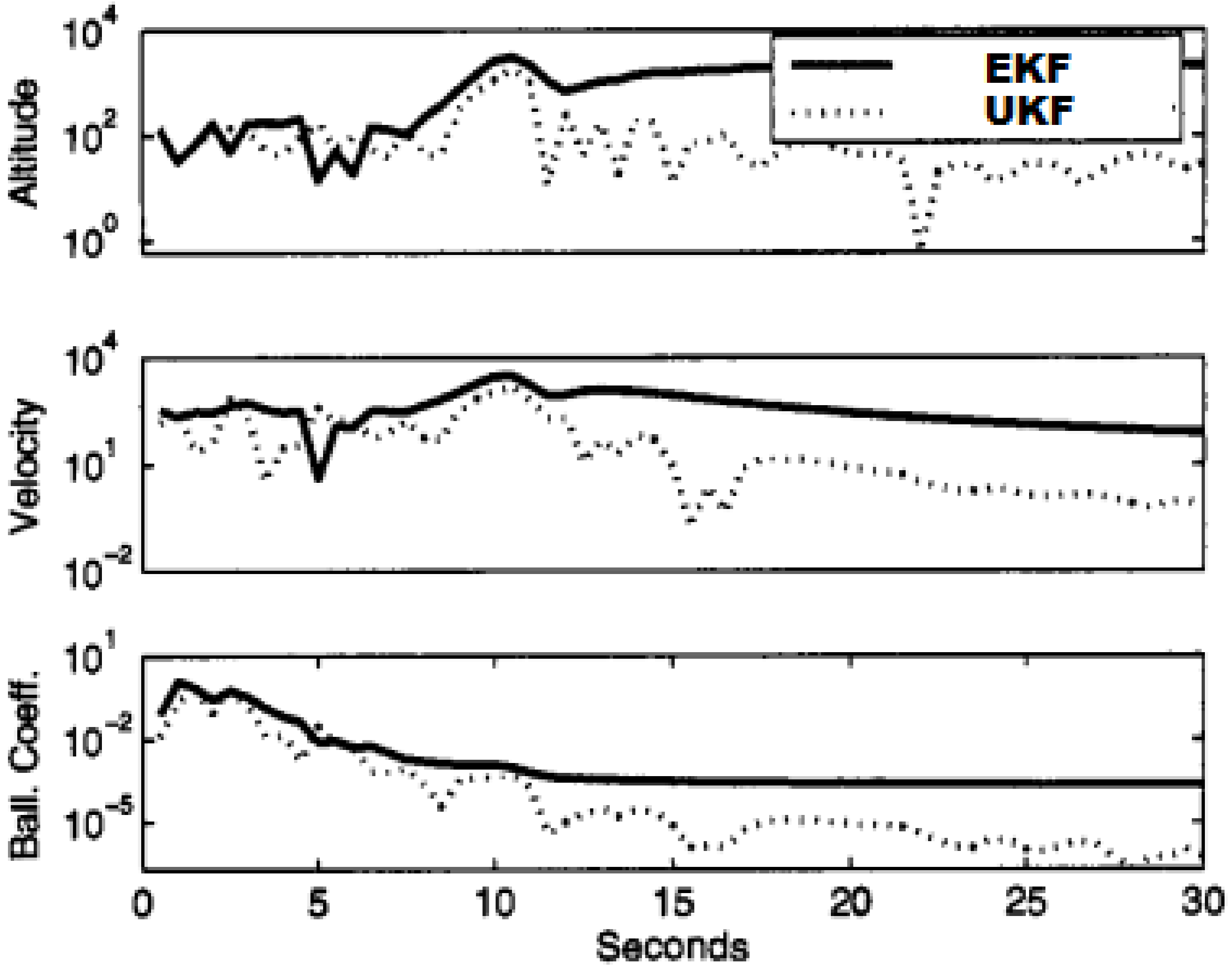
$$y(t_k) = \sqrt{M^2 + \left[x_1(t_k) - a\right]^2} + v_k$$



Example V

Free falling of a particle:

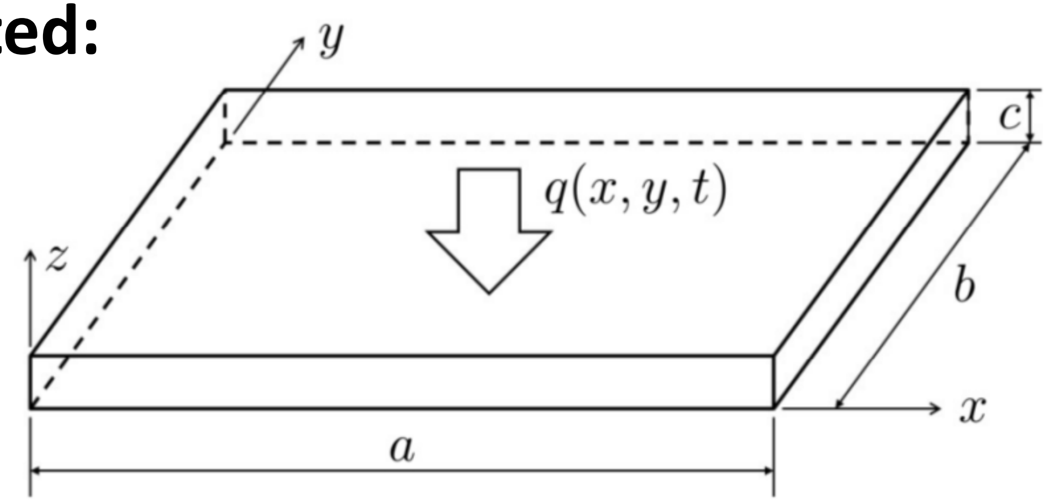




Example VI – Introduction

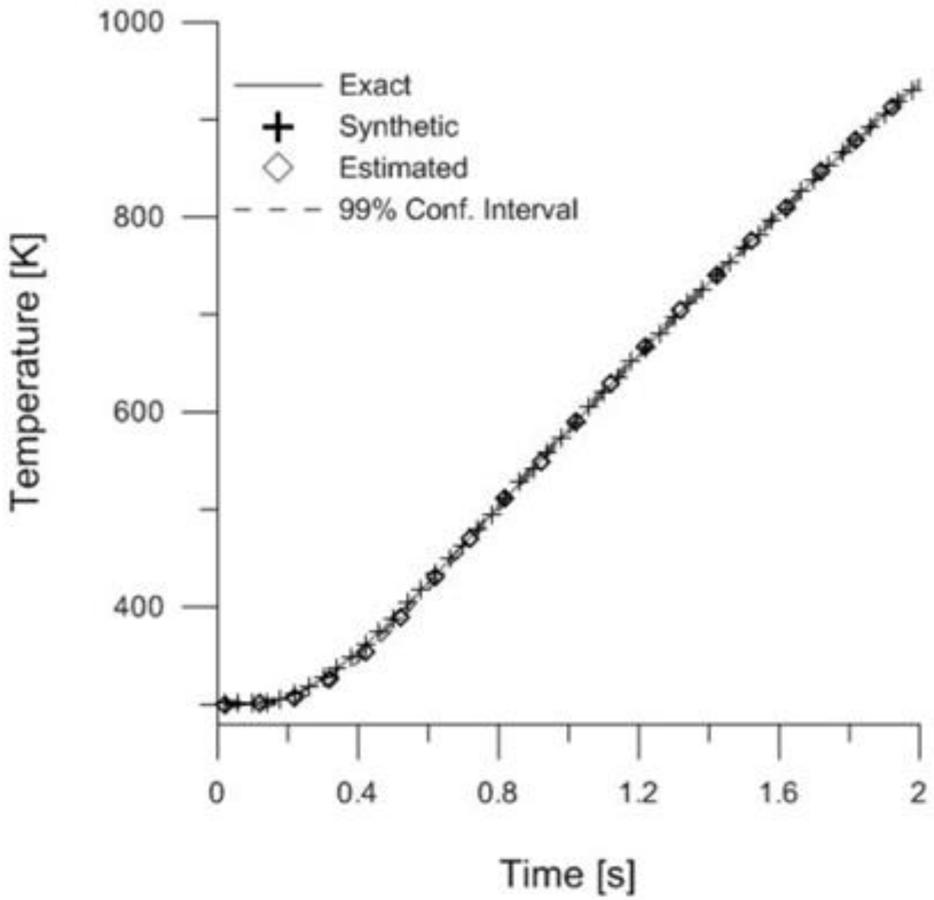
Heat Flux Identification problem – revisited:

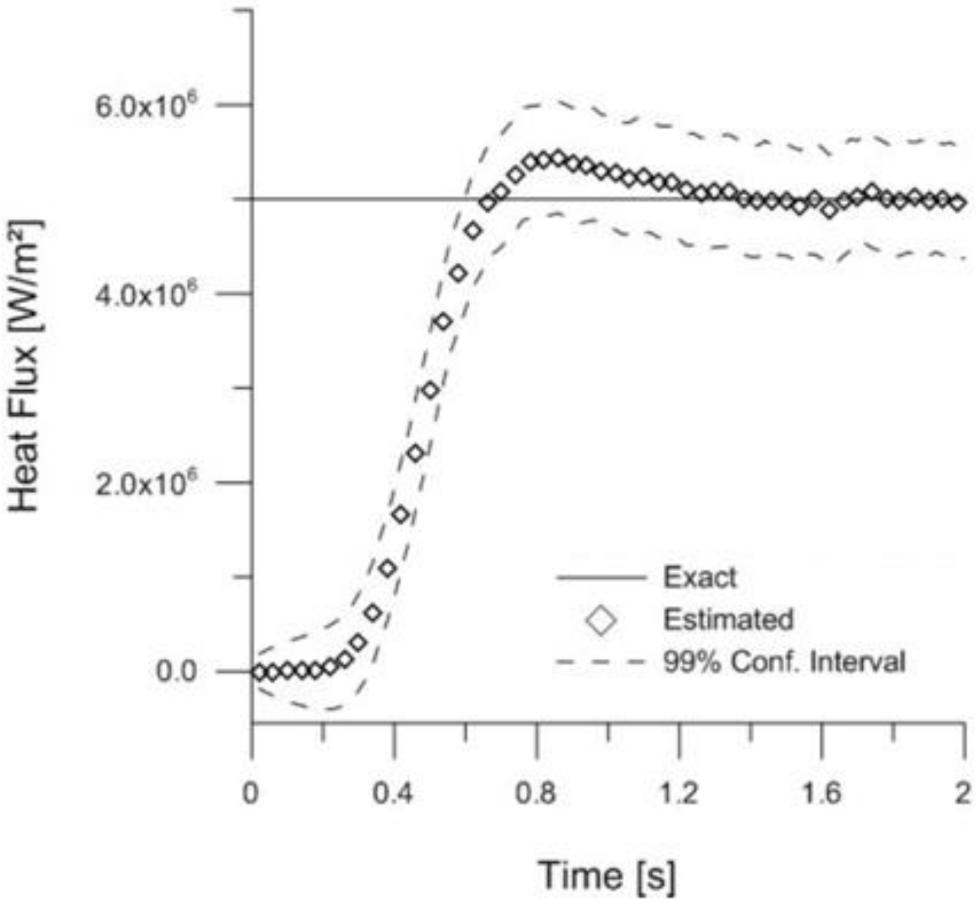
- Same model as example III;
- Estimation of heat flux at z=0;
- Temperature measurements at z=c;
- 3D, Nonlinear;
- Heat flux application from t=0;
- Use Complete Model for Inverse Problem solution;



Example VI – Results

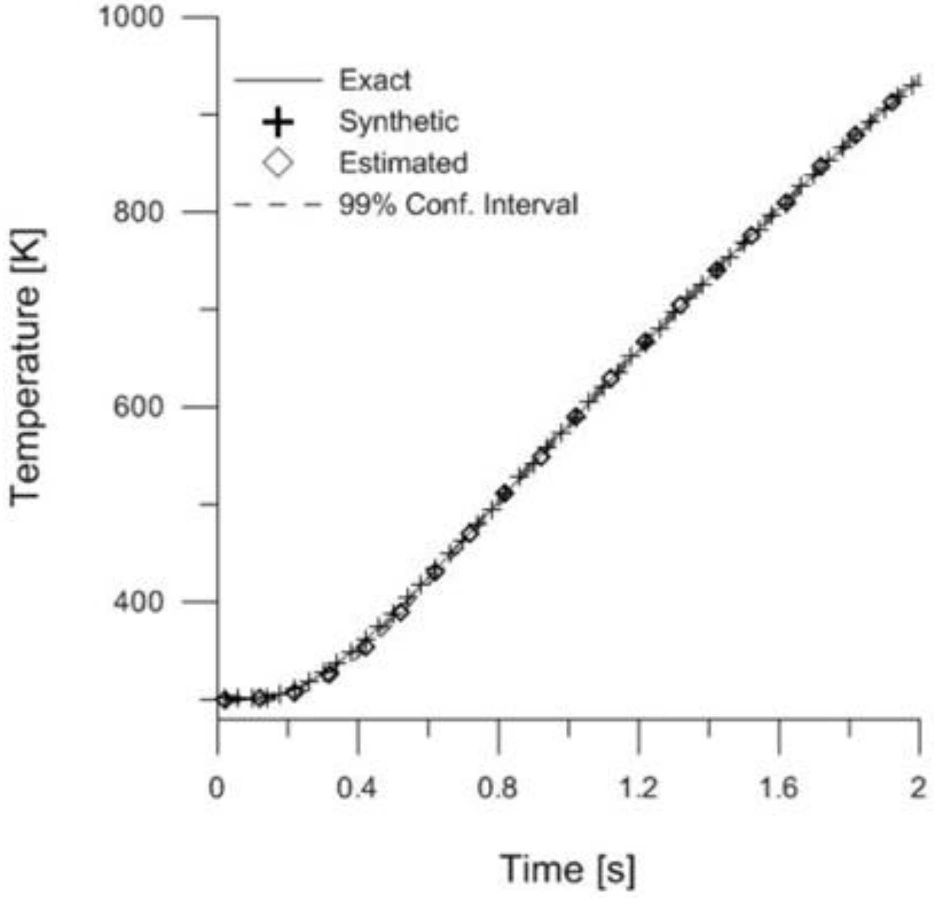
Hot Spot #2: Temperature and heat flux estimates

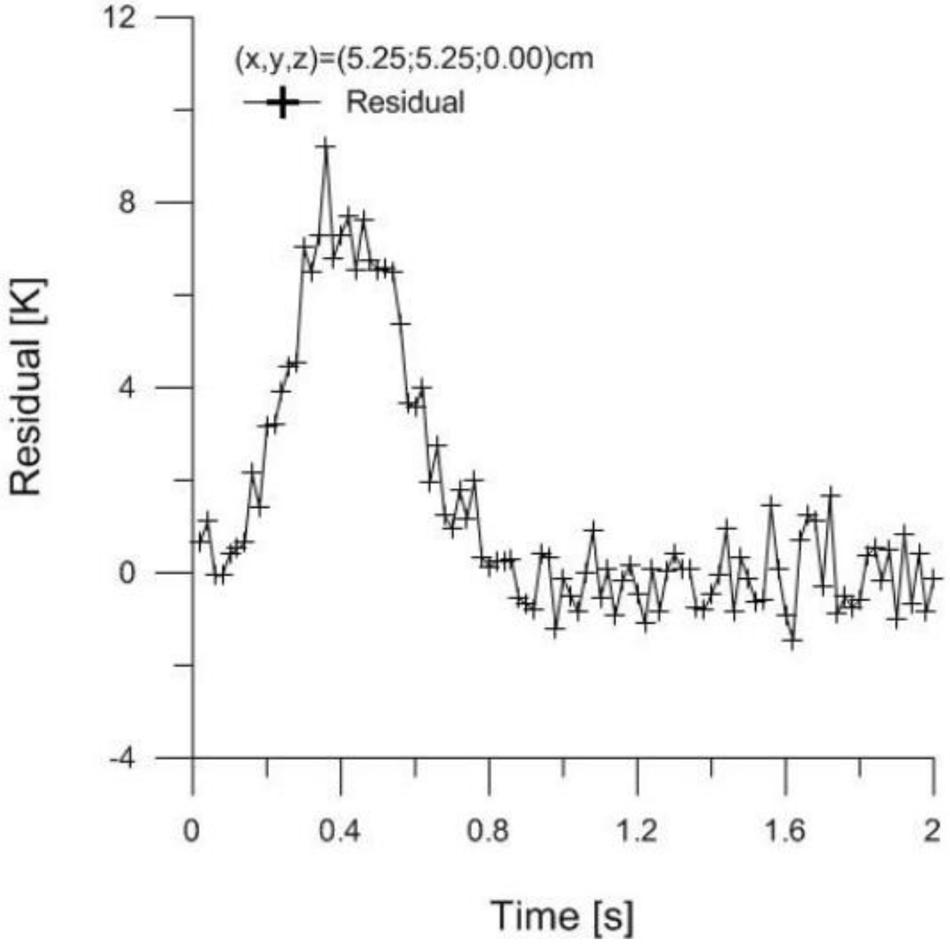




Example VI – Results

Hot Spot #2: Temperature estimates and residuals





Example VI – Results

Computational time:

- Physical time: 2.0 s;
- Kalman filter: N/A;
- **UKF: 115200 s (approx. 32 h);**

UKF and use of the Complete Model:

- Inverse Analysis is made possible;
- Computational time is excessively large.

Overall Conclusions

Kalman filtering techniques

- Wide range of applications for sequential estimation;
- Much more variations available;

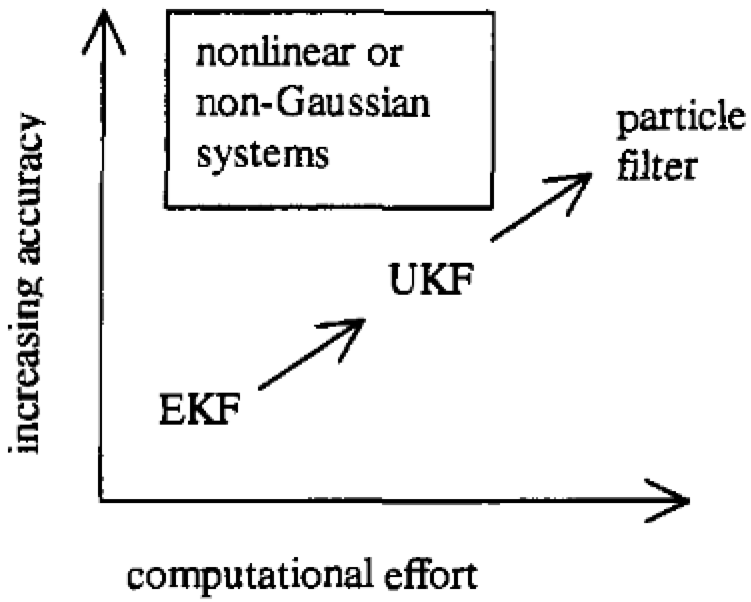
Classical Kalman Filter:

- Analytical solution for linear and Gaussian problems;
- Stable, robust, easy to program;

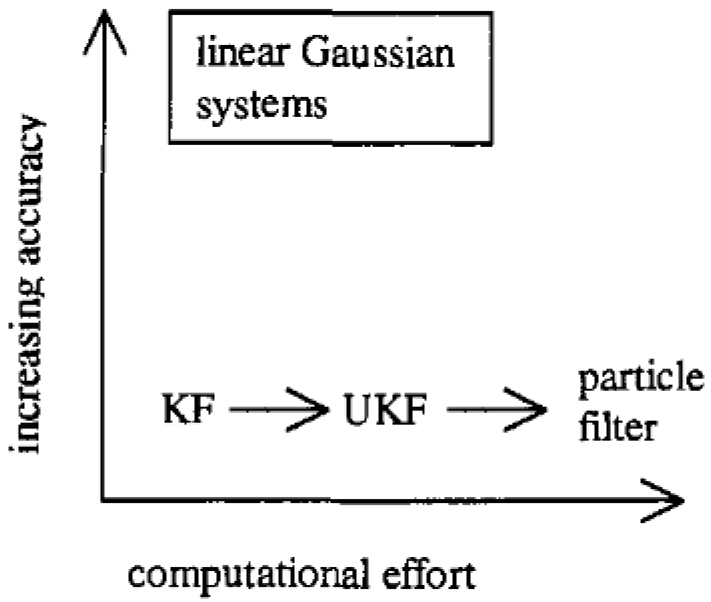
Extended and Unscented Kalman Filter:

- Reliable algorithms for nonlinear and/or non Gaussian problems;
- Increasing degrees of accuracy and complexity.

Overall Conclusions



(a) The above figure depicts the increasing computational effort and increasing accuracy that is obtained by going from an EKF to a UKF to a particle filter. This applies to systems that are nonlinear or non-Gaussian.



(b) The above figure depicts the fact that the Kalman filter is optimal for linear Gaussian systems. Going from a Kalman filter to a UKF to a particle filter will increase computational effort but will not improve estimation accuracy.

SIMON D., *Optimal State Estimation*, John Wiley & Sons, 2006.