

“Experimental Data Processing”

Topic 3

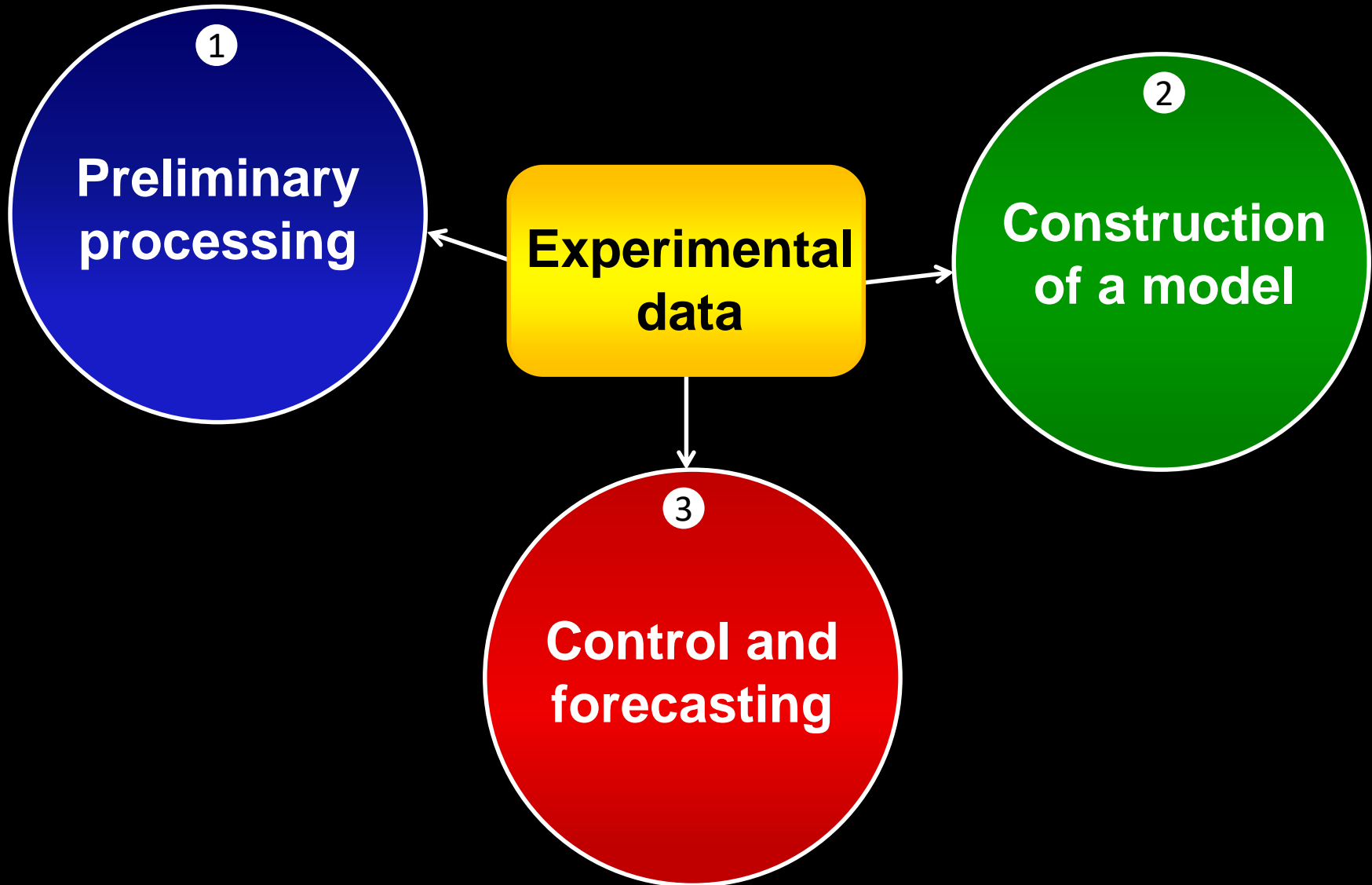
"Optimal approximation at state space"

Tatiana Podladchikova

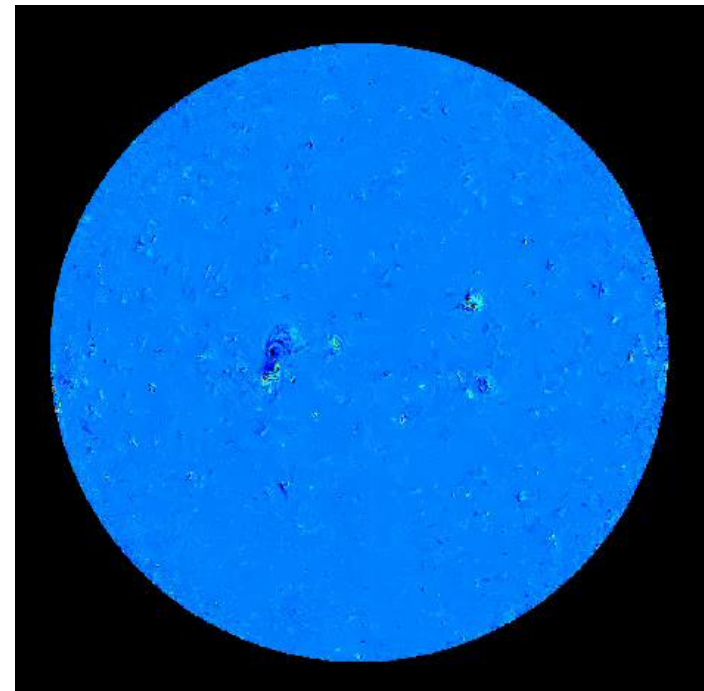
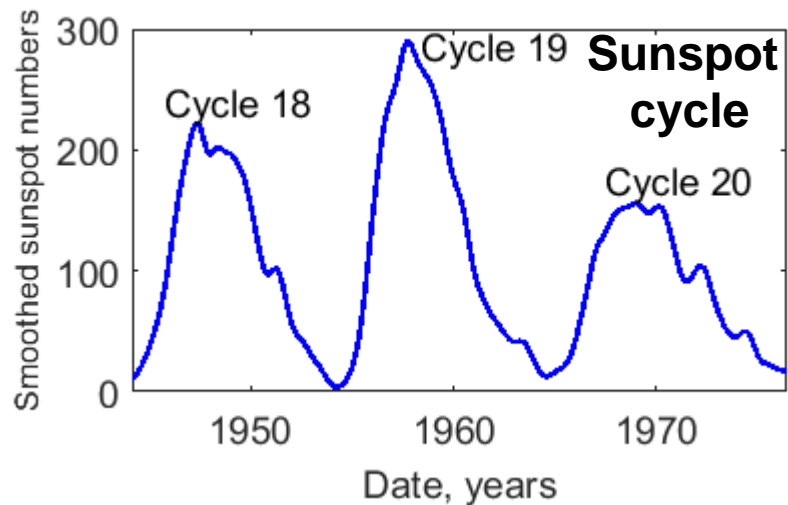
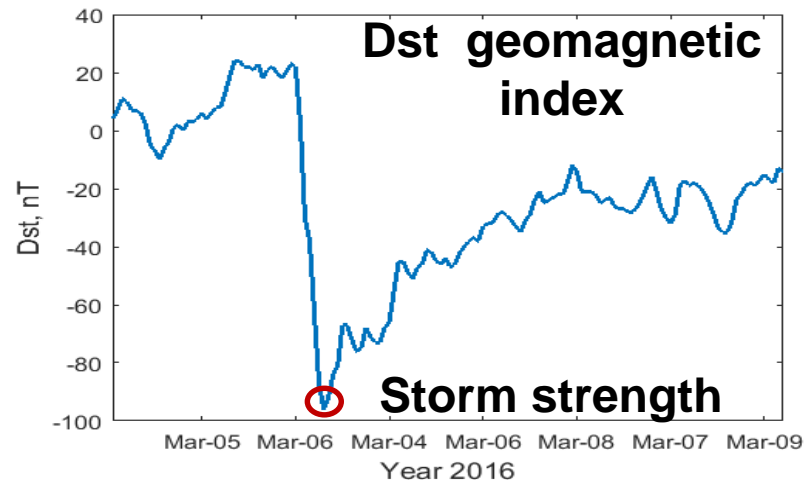
Term 1B, October 2019

t.podladchikova@skoltech.ru

Traditional approach to estimation and forecasting



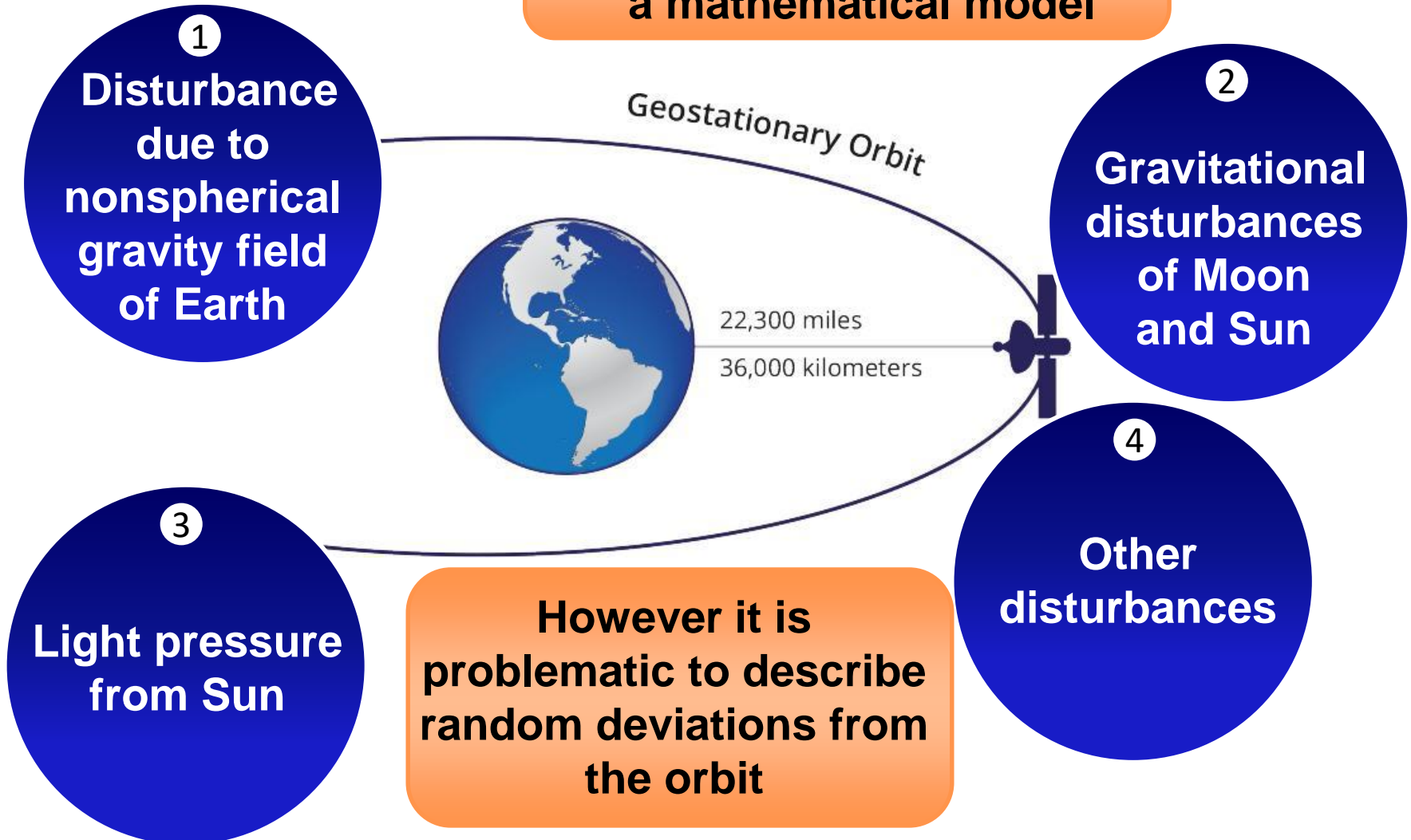
Creation of a mathematical model of insufficiently studied processes is quite problematic



**Extreme ultraviolet
coronal wave
December 7, 2007**


Forces that affect motion of a geostationary satellite

Motion of controlled objects is usually described by a mathematical model



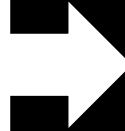
Application area of quasi-optimal methods

**Mathematical model
requires prior assumptions
about a process**



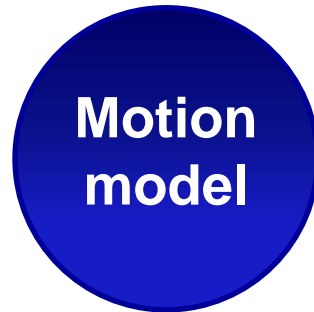
**These assumptions
may significantly distort
estimation output**

**Quasi-optimal methods
do not need any prior
assumptions that may
distort a process**

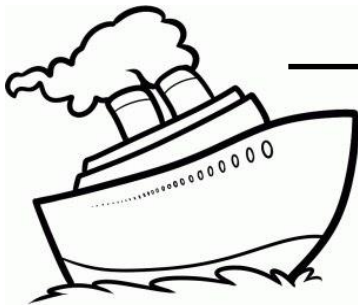


**Thus they can extract
hidden regularities for
long-term forecasting of
complicated processes**

From Gauss to Kalman



$$x_i = x_{i-1} + V_{i-1}T + \frac{a_i T^2}{2}$$
$$V_i = V_{i-1} + a_i T$$



V_i

Unintentional maneuver can be described by random acceleration a_i

ship pitching or undercurrents

x

Classical least – square method provides estimations of constant parameters

A. Legendre, 1806
J. Gauss, 1809

Development

Kalman filter provides estimations of variable parameters x_i, V_i

R. Kalman, 1960

State equation

Motion
model



$$x_i = x_{i-1} + V_{i-1}T + \frac{a_i T^2}{2}$$
$$V_i = V_{i-1} + a_i T$$

$$X_i = \begin{bmatrix} x_i \\ V_i \end{bmatrix}$$

State
vector

It contains full
information about the
state of system at time i

State equation

Motion
model

$$\begin{aligned}x_i &= x_{i-1} + V_{i-1}T + \frac{a_i T^2}{2} \\V_i &= V_{i-1} + a_i T\end{aligned}$$

$$X_i = \begin{bmatrix} x_i \\ V_i \end{bmatrix}$$

State
vector

It contains full
information about the
state of system at time i

State
equation

$$X_i = \Phi X_{i-1} + G a_i$$

$$\Phi = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$$

Transition
matrix

$$G = \begin{bmatrix} T^2/2 \\ T \end{bmatrix}$$

Input
matrix

Measurement equation

Measurements
of coordinate x_i
with error η_i

$$z_i = x_i + \eta_i$$

Measurement
equation

$$z_i = HX_i + \eta_i$$

$$X_i = \begin{bmatrix} x_i \\ v_i \end{bmatrix}$$

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

Observation
matrix

State space model

State
equation

$$X_i = \Phi_{i,i-1}X_{i-1} + w_i$$

Stochastic
model

X_i
State
vector

$\Phi_{i,i-1}$
Transition matrix,
that relates states
 X_i and X_{i-1}

w_i
State noise
describing model
errors with
covariance matrix Q_i

Measurement
equation

$$z_i = H_iX_i + \eta_i$$

z_i
Measurements

H_i
Observation matrix,
that relates state X_i
with measurements z_i

η_i
Measurement noise
with the covariance
matrix R_i

State space model

State
equation

$$X_i = \Phi_{i,i-1}X_{i-1} + w_i$$

Measurement
equation

$$z_i = H_iX_i + \eta_i$$

w_i

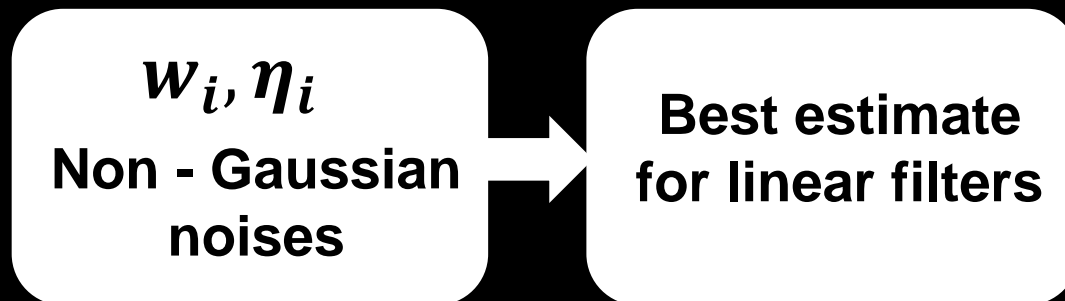
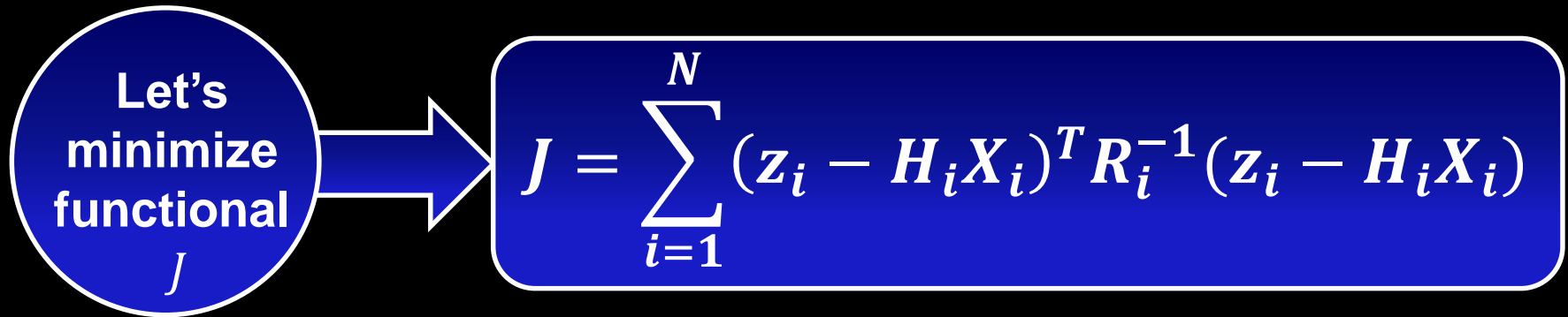
Noise intrinsic to
the process itself
that should not
be filtered

State space
model separates
noises in
contrast to linear
regression

η_i

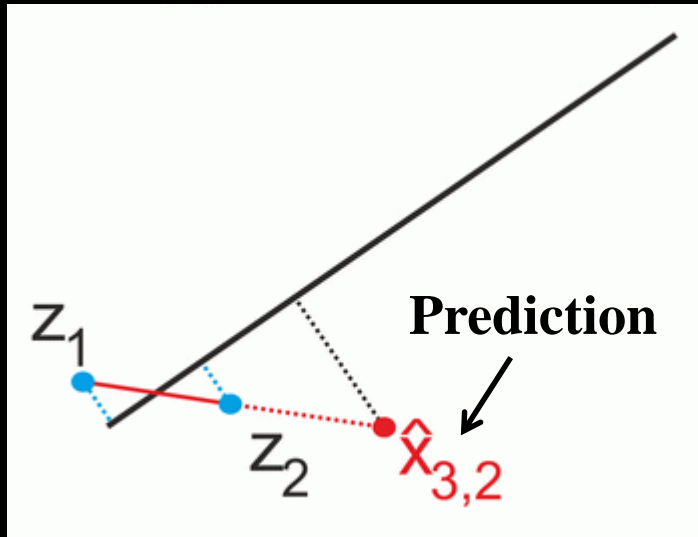
Measurement
noise that
should be
filtered

Kalman filter estimate from Least-Squares method

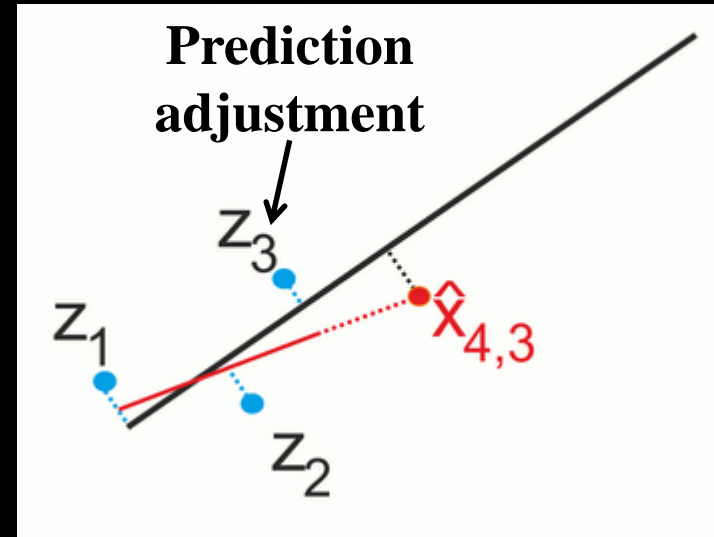


Recurrent algorithm of Kalman filter

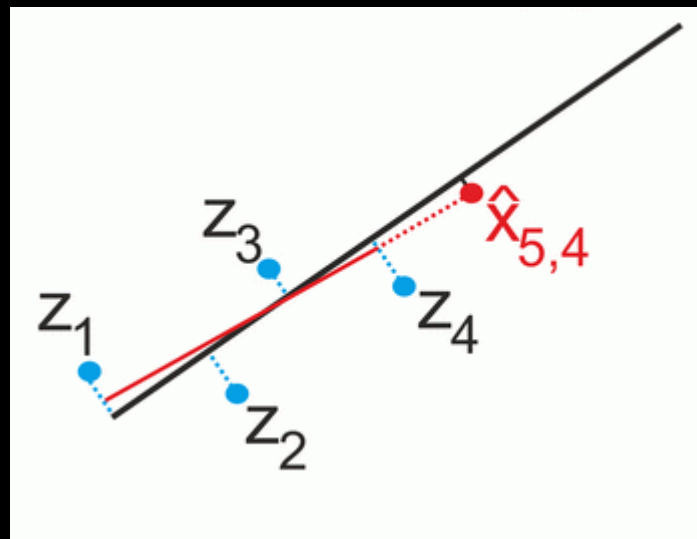
2 measurements



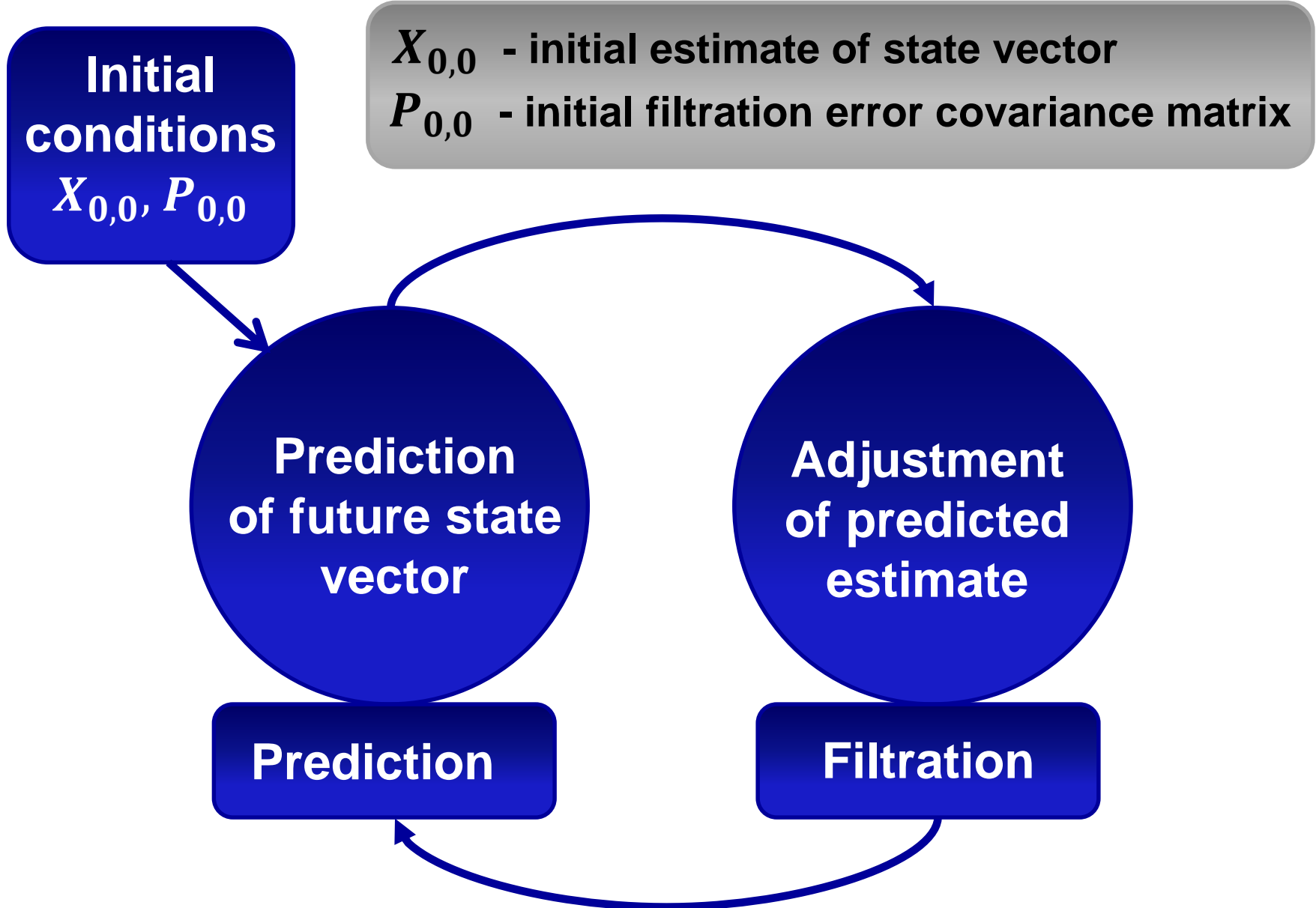
3 measurements



4 measurements



Recurrent algorithm of Kalman filter



Recurrent algorithm of Kalman filter

① Prediction (extrapolation)

Prediction of state vector at time i using $i - 1$ measurements

$$X_{i,i-1} = \Phi_{i,i-1} X_{i-1,i-1}$$

Prediction error covariance matrix

$$P_{i,i-1} = \Phi_{i,i-1} P_{i-1,i-1} \Phi_{i,i-1}^T + Q_i$$

$$P_{i,i-1} = E[(X_i - X_{i,i-1})(X_i - X_{i,i-1})^T]$$

$X_{i,i-1}$

First subscript i
denotes time on which
the prediction is made

Second subscript $i - 1$
represents the number of
measurements to get $X_{i,i-1}$

Recurrent algorithm of Kalman filter

② Filtration

Adjustment of predicted estimate

Improved estimate by incorporating a new measurement

$$X_{i,i} = X_{i,i-1} + K_i(z_i - HX_{i,i-1})$$

Residual

Filter gain, weight of residual

$$K_i = P_{i,i-1}H_i^T(H_iP_{i,i-1}H_i^T + R_i)^{-1}$$

Filtration error covariance matrix

$$P_{i,i} = (I - K_iH_i)P_{i,i-1}$$

$$P_{i,i} = E[(X_i - X_{i,i})(X_i - X_{i,i})^T]$$

**Classical Least-Squares method (LSM)
is particular case of Kalman filter**



**Dynamical
model
is deterministic.
Covariance matrix
of state noise w
 $Q = 0$**

**The Kalman filter
solution is equivalent
to that of LSM**



**However recurrent form of
Kalman filter solution has great
advantage for implementation**

1

**Nonlinear
dynamical
model**

2

**Nonlinear
relation between
state and
measurement
vector**

3

**Correlated
state noise**

4

**Correlated
measurement
noise**

5

**Correlation
between state
and
measurement
noise**

6

**Biased
state noise
and/or
measurement
noise**

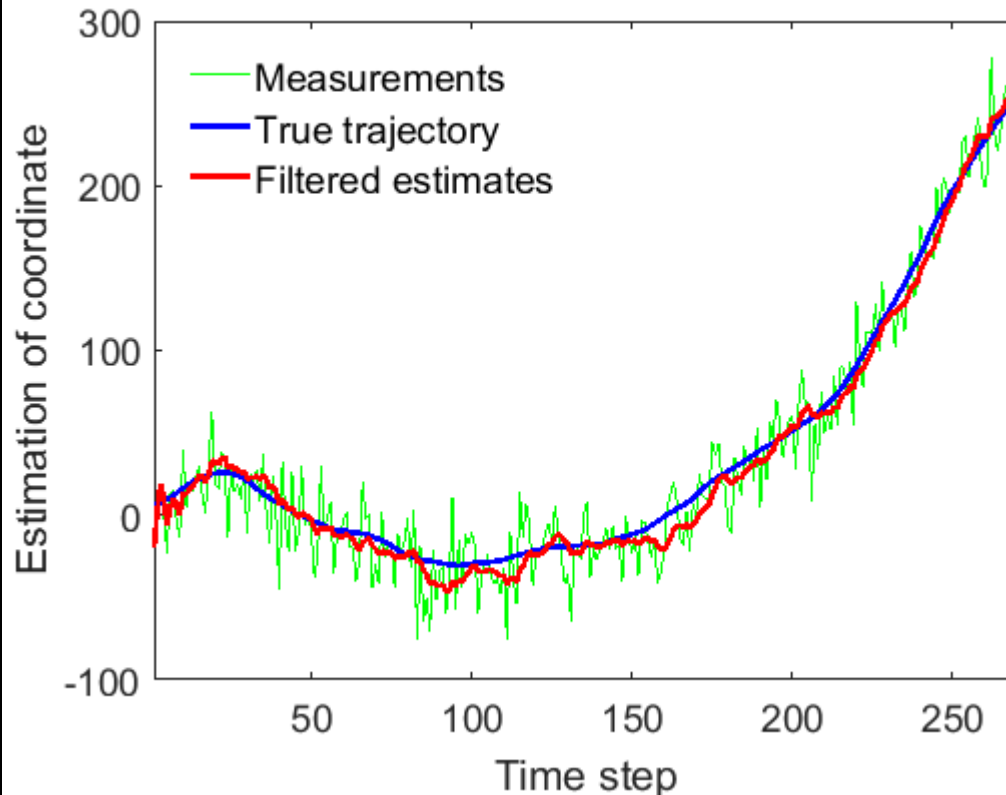
**Kalman filter
modifications**

Tracking moving object using Kalman filter

Stochastic model

Motion
model

$$x_i = x_{i-1} + V_i T + \frac{a_i T^2}{2}$$
$$V_i = V_{i-1} + a_i T$$



$$P_{0,0}^{-1} = \infty$$
$$\sigma_a^2 = 0.04$$
$$\sigma_\eta^2 = 400$$

Tracking moving object using Kalman filter

Stochastic model

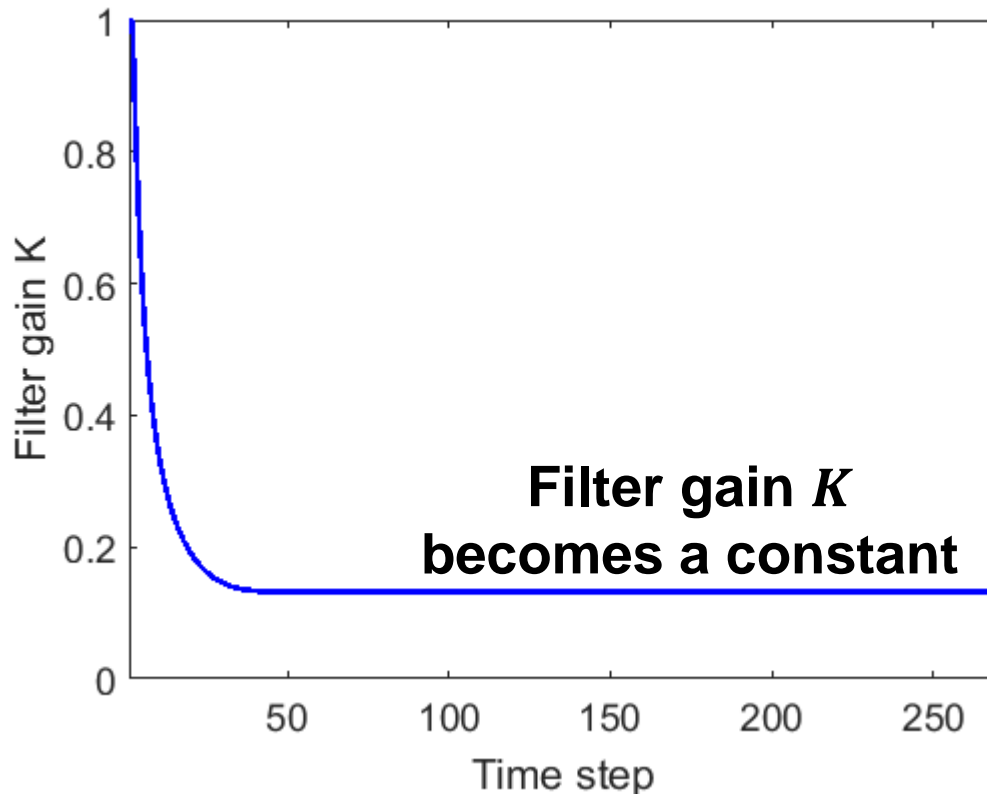
Motion
model

$$x_i = x_{i-1} + V_i T + \frac{a_i T^2}{2}$$
$$V_i = V_{i-1} + a_i T$$

Kalman filter
becomes stationary

After that there is
no increase of
estimation accuracy

Measurements
always adjust
prediction

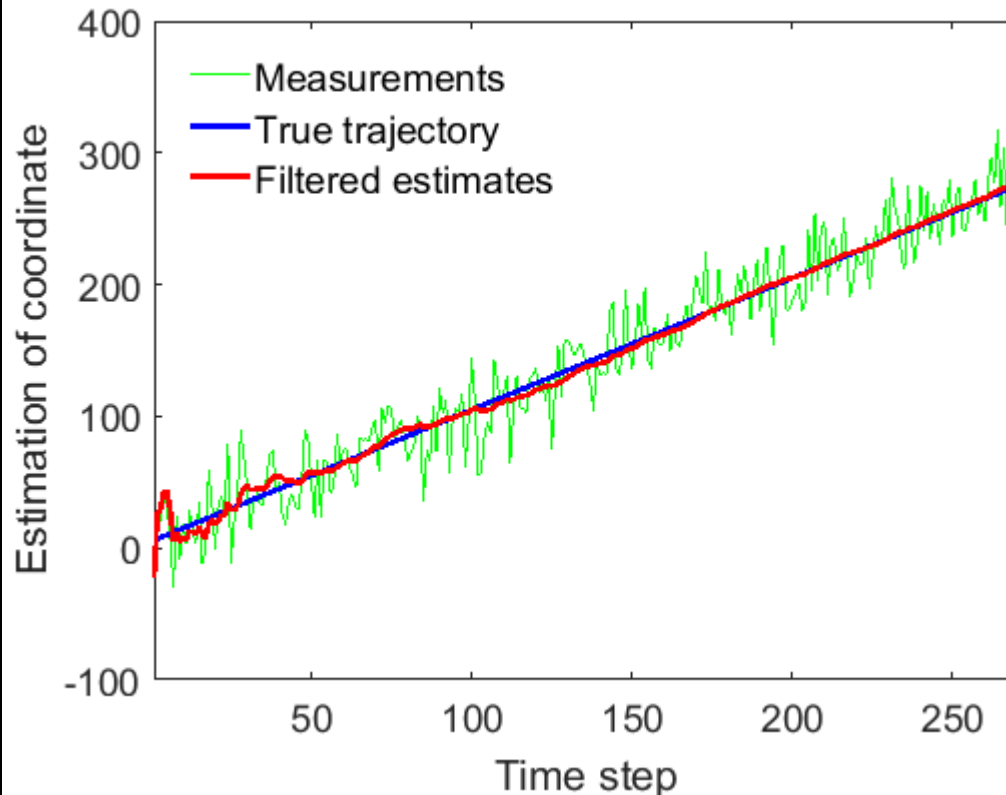


Tracking moving object using Kalman filter

Deterministic model

Uniform
motion

$$x_i = x_{i-1} + V_i T$$
$$V_i = V_{i-1}$$



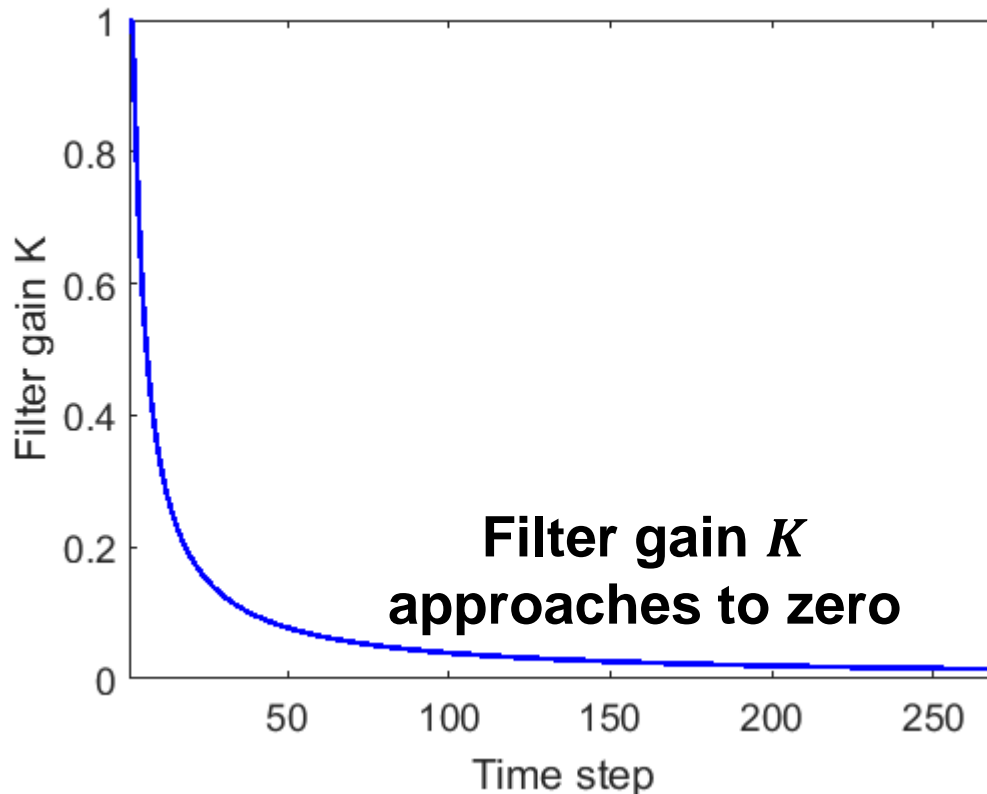
$$P_{0,0}^{-1} = \infty$$
$$\sigma_a^2 = 0$$
$$\sigma_\eta^2 = 400$$

Tracking moving object using Kalman filter

Deterministic model

Uniform
motion

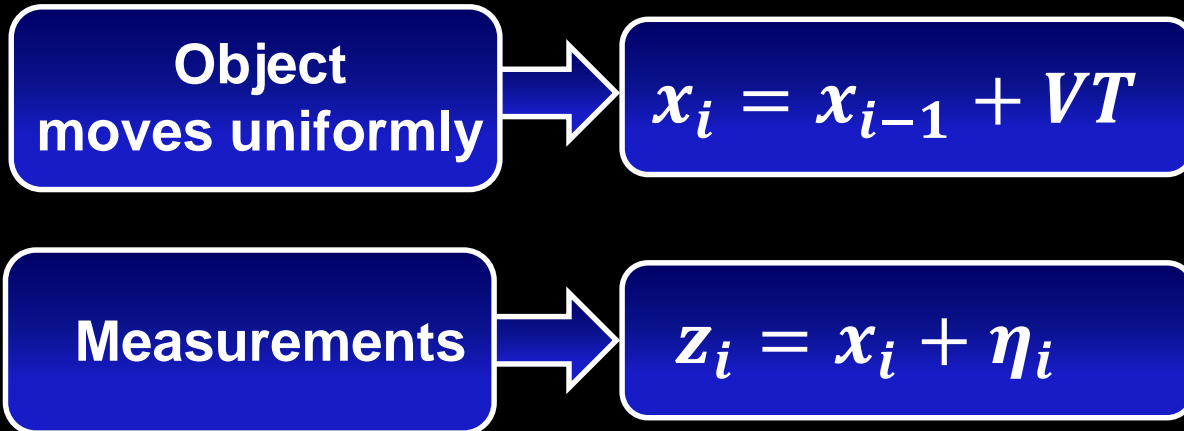
$$x_i = x_{i-1} + V_i T$$
$$V_i = V_{i-1}$$



High estimation
accuracy achieved

Filter switches
off from
measurements

Alpha-beta filter – simplified case of Kalman filter



Let's use these parameters in Kalman filter algorithm

$$X_i = \begin{bmatrix} x_i \\ V_i \end{bmatrix} \quad \Phi = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \quad H = \begin{bmatrix} 1 & 0 \end{bmatrix}$$
$$Q = 0 \quad P_{0,0}^{-1} = 0$$

Alpha-beta filter – simplified case of Kalman filter

Predicted estimate

$$x_{i,i-1} = x_{i-1,i-1} + V_{i-1,i-1}T$$

$$V_{i,i-1} = V_{i-1,i-1}$$

Filtered estimate

$$x_{i,i} = x_{i,i-1} + \alpha(z_i - x_{i,i-1})$$

$$V_{i,i} = V_{i,i-1} + \beta(z_i - x_{i,i-1})$$

$$\alpha = \frac{2(2i-1)}{i(i+1)}$$

$$\beta = \frac{6}{i(i+1)T}$$

Why does
distance between
true trajectory
and estimation
increase?



Divergence. Errors
monotonously
increase

Alpha-beta filter – simplified case of Kalman filter

Filtered estimate

$$x_{i,i} = x_{i,i-1} + \alpha(z_i - x_{i,i-1})$$

$$V_{i,i} = V_{i,i-1} + \beta(z_i - x_{i,i-1})$$

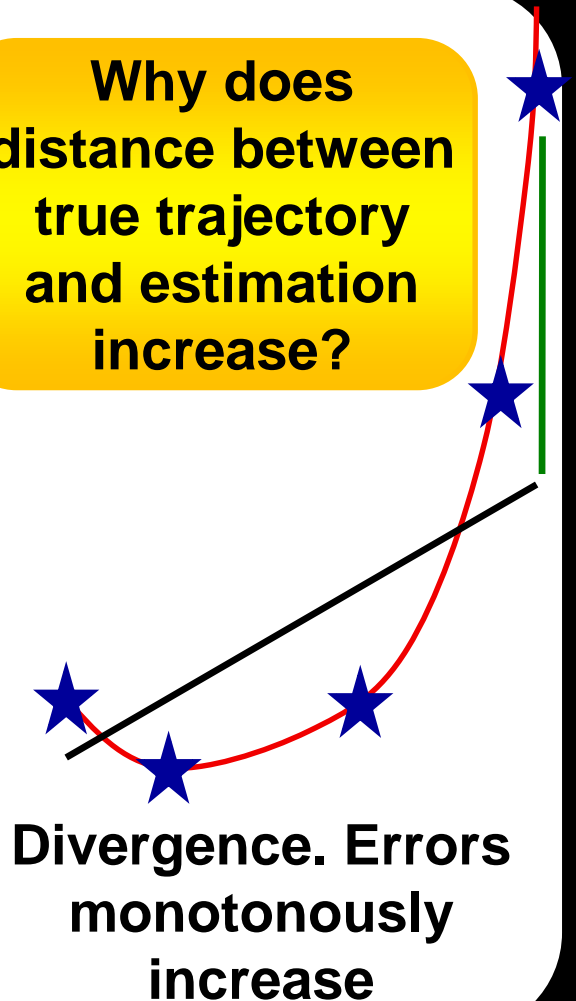
$$\alpha = \frac{2(2i - 1)}{i(i + 1)}$$

$$\beta = \frac{6}{i(i + 1)T}$$

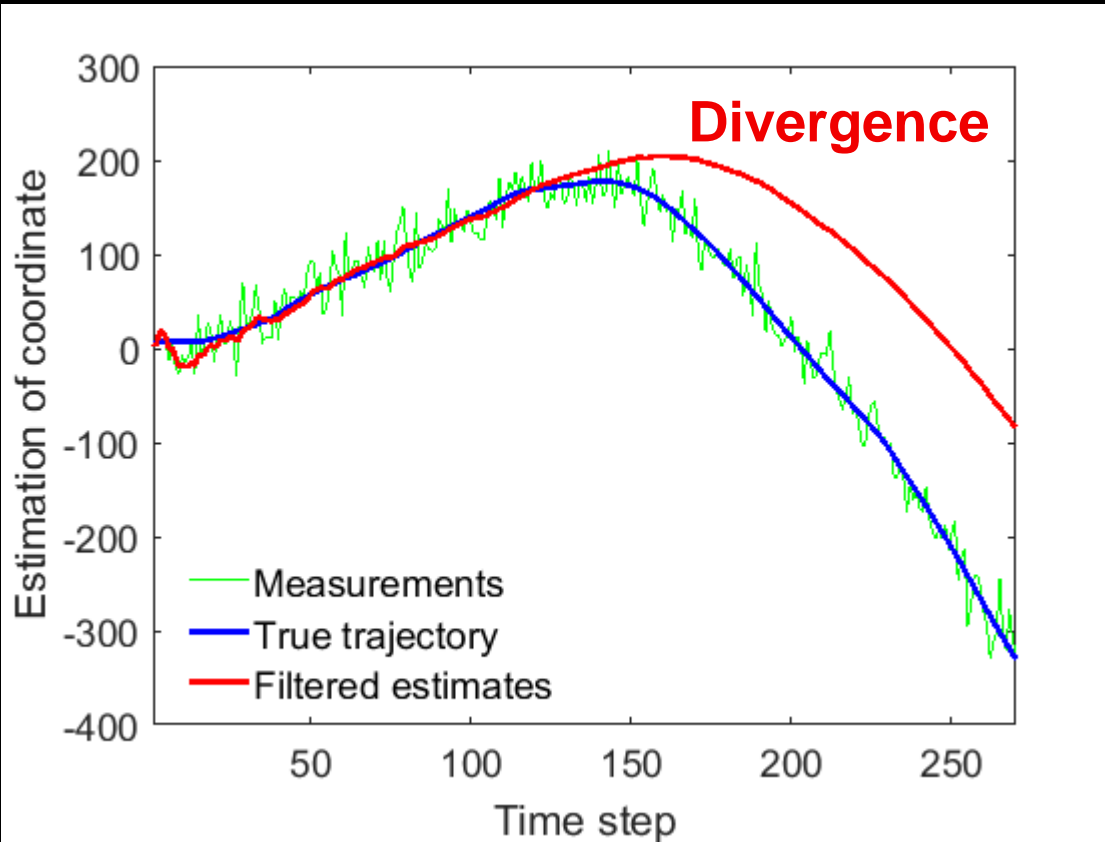
With increase
of i coefficients
 $\alpha, \beta \rightarrow 0$

Filter switches
off from
measurements

Why does
distance between
true trajectory
and estimation
increase?



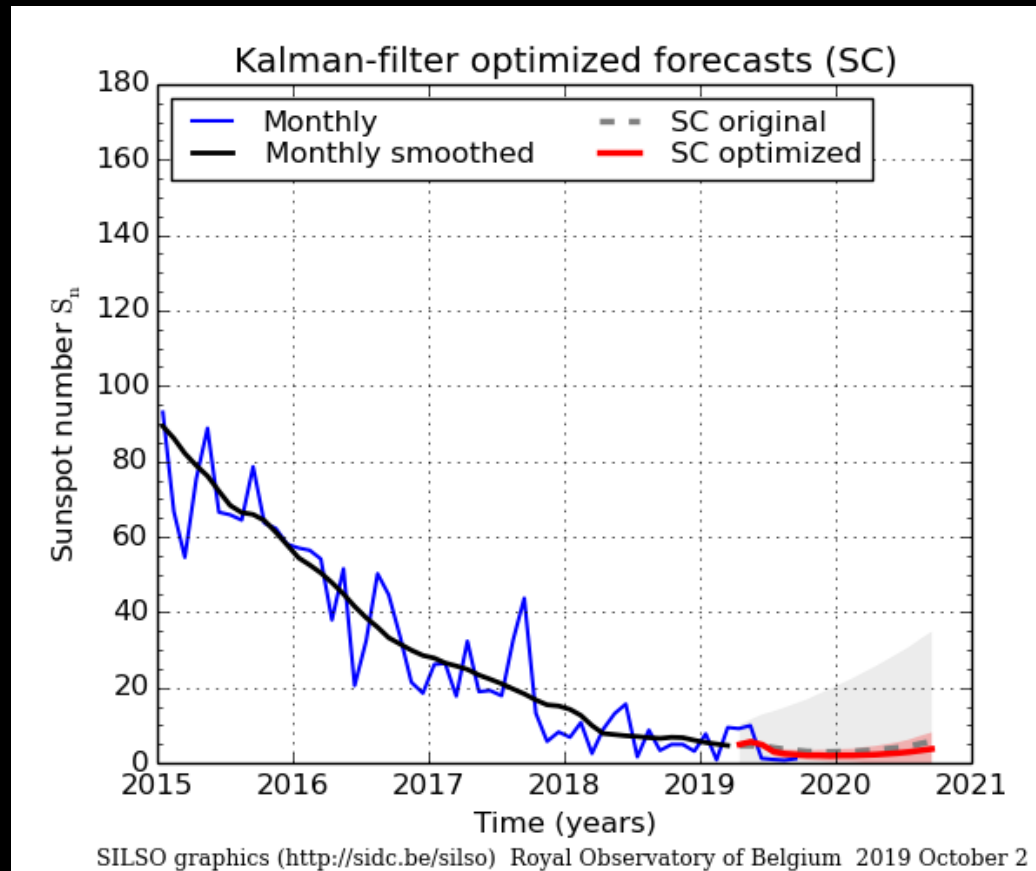
What happens if we use deterministic model, but in fact it is stochastic model?



Filter gain K
approaches to zero
for deterministic
model

Filter diverges
as it switches off
from measurements

Kalman filter to forecast sunspot number 12 months ahead

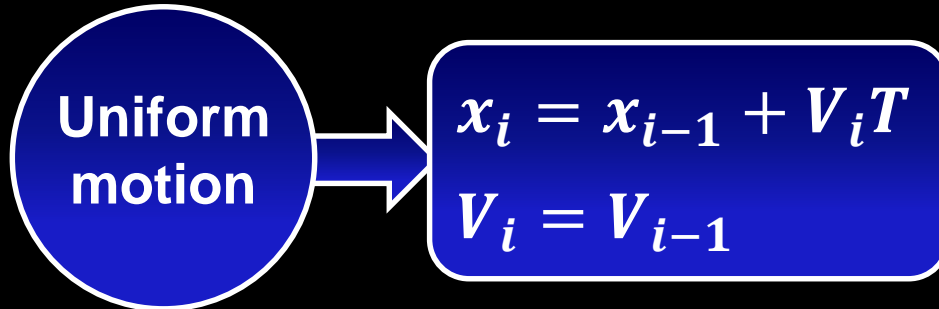


Extrapolation 12 steps ahead

$$X_{12,1} = \Phi_{12,1} X_{1,1}$$

$$\Phi_{12,1} = \Phi_{12,11} \Phi_{11,10} + \cdots \Phi_{3,2} \Phi_{2,1}$$

Non-observable system. Example 1



**Measurements of only
velocity V_i are available**

$$z_i = V_i + \eta_i$$

**Measurements of coordinate
 x_i are not available**

Non-observable system. Example 1

Uniform
motion

$$\begin{aligned}x_i &= x_{i-1} + V_i T \\ V_i &= V_{i-1}\end{aligned}$$

Measurements of only
velocity V_i are available

$$z_i = V_i + \eta_i$$

Measurements of coordinate
 x_i are not available

Let's present the system at state space

State
equation

$$X_i = \Phi X_{i-1}$$

Measurement
equation

$$z_i = H X_i + \eta_i$$

$$X_i = \begin{bmatrix} x_i \\ V_i \end{bmatrix}$$

State
vector

$$\Phi = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$$

Transition
matrix

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

Observation
matrix

Non-observable system. Example 1

Uniform
motion

$$\begin{aligned}x_i &= x_{i-1} + V_i T \\ V_i &= V_{i-1}\end{aligned}$$

Measurements of only
velocity V_i are available

$$z_i = V_i + \eta_i$$

Measurements of coordinate
 x_i are not available

Let's present the system at state space

State
equation

$$X_i = \Phi X_{i-1}$$

Measurement
equation

$$z_i = H X_i + \eta_i$$

$$X_i = \begin{bmatrix} x_i \\ V_i \end{bmatrix}$$

State
vector

Is it possible to estimate
coordinate x_i using Kalman filter?

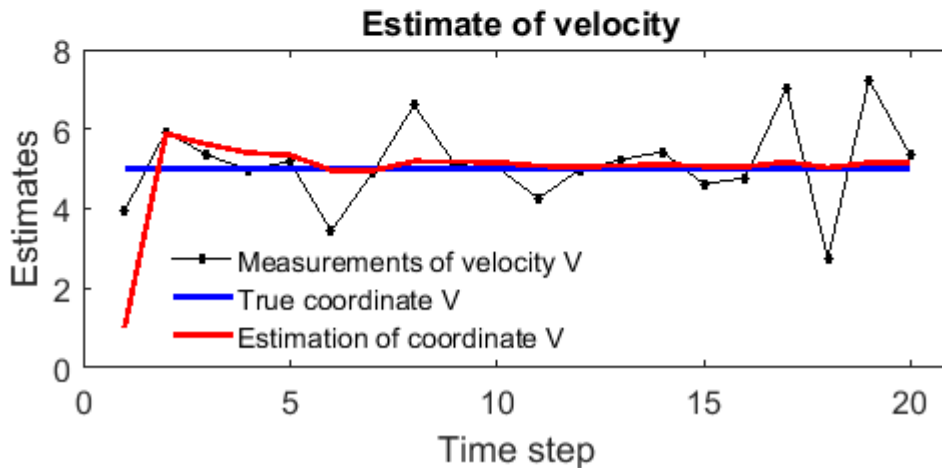
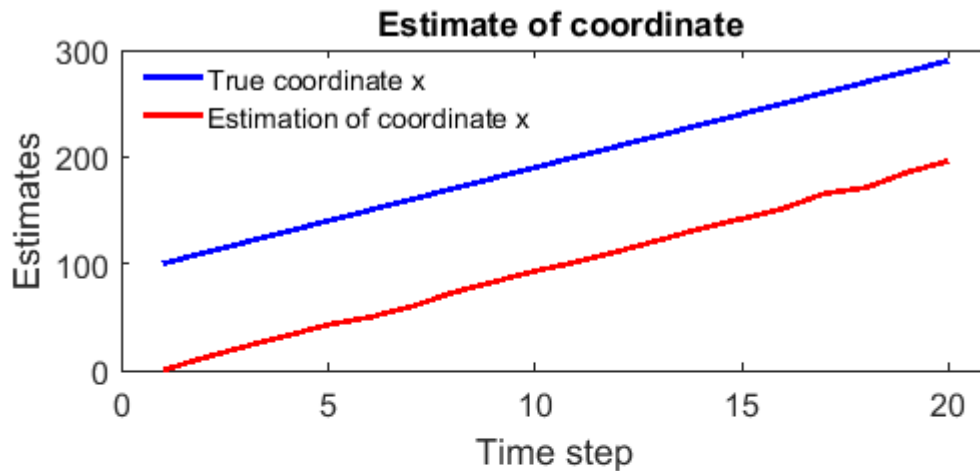
$$\Phi = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$$

Transition
matrix

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

Observation
matrix

Non-observable system. Example 1



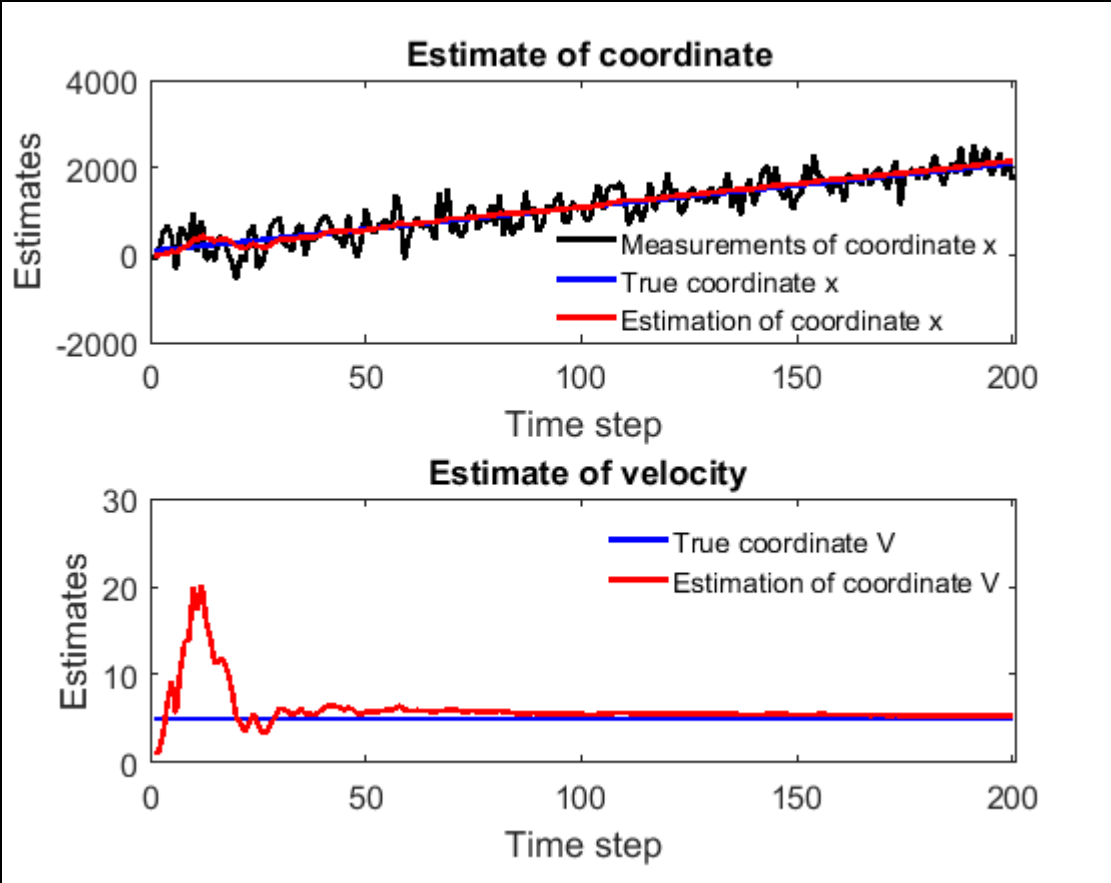
Coordinate x_i cannot be adjusted by measurements of V_i

Kalman filter minimizes the variance of estimation error of velocity V , but not the coordinate x

The term “optimality” is applicable only for observable components

The initial error x_0 is kept during all the filtration interval

Non-observable system. Example 1

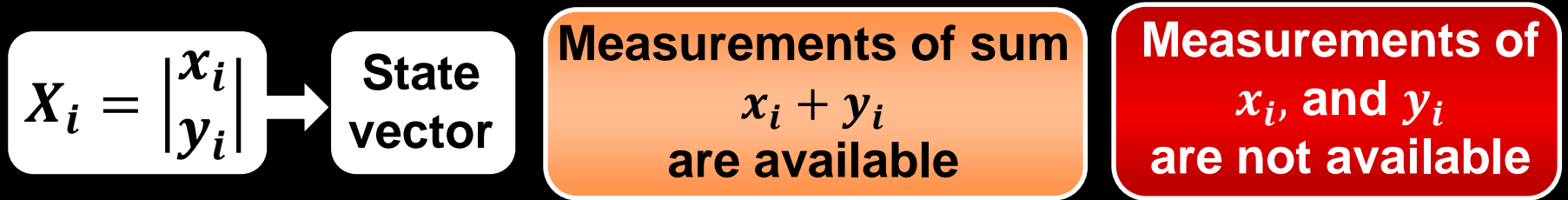


**Measurements
of only coordinate x_i
are available**

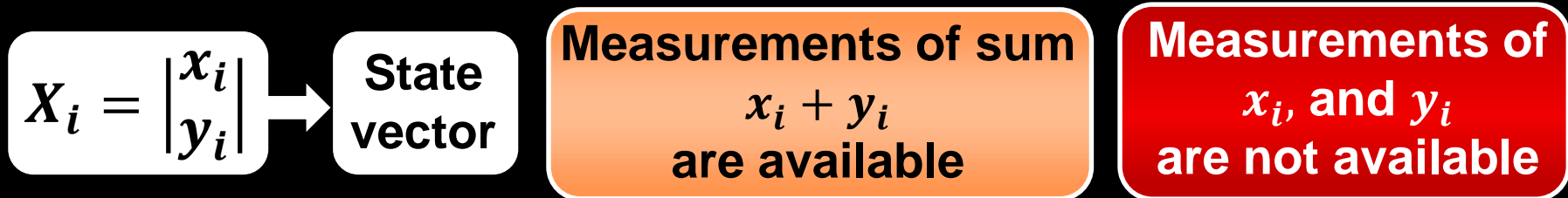
System is observable

**Kalman filter provides
estimation of full
state vector X_i**

Non-observable system. Example 2



Non-observable system. Example 2



The system at state space



Is it possible to estimate state vector X_i ?

To apply Kalman filter we need to analyze observability of a system

The observability is determined only by pair of matrices Φ and H .



$$\text{rank}[H^T \ \Phi^T H^T \ (\Phi^T)^2 H^T \ \dots \ (\Phi^T)^{n-1} H^T] = n$$

For stationary system

n – dimension of state vector

To apply Kalman filter we need to analyze observability of a system

The observability is determined only by pair of matrices Φ and H .



$$\text{rank}[H^T \Phi^T H^T (\Phi^T)^2 H^T \dots (\Phi^T)^{n-1} H^T] = n$$

For stationary system

n – dimension of state vector

$$\text{rank}[H^T \Phi^T H^T (\Phi^T)^2 H^T \dots (\Phi^T)^{n-1} H^T] = q < n$$

Partial observability

$$\frac{q}{n}$$



Observability degree

To apply Kalman filter we need to analyze observability of a system

$$\text{rank}[H^T \Phi^T H^T (\Phi^T)^2 H^T \dots (\Phi^T)^{n-1} H^T] = n$$

Analysis of system observability for example 1

$$\Phi = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$$

Transition matrix

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

Observation matrix

$$X_i = \begin{bmatrix} x_i \\ V_i \end{bmatrix}$$

Dimension of state vector $n = 2$

$$\text{rank}[H^T \Phi^T H^T] = \text{rank} \left[\begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right] = \text{rank} \left[\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \right] = 1$$

System is partly observable

Only one component is observable

To apply Kalman filter we need to analyze observability of a system

Observability Gramian W
for non-stationary system



$$W = \sum_{i=1}^n \Phi_{i,n}^T H_i^T H_i \Phi_{i,n} > 0$$

Positive-
definite matrix

$\Phi_{i,n}$ is inverse matrix to transition matrix $\Phi_{n,i}$
$$\Phi_{n,i} = \Phi_{n,n-1} \cdot \Phi_{n-1,n-2} \cdots \cdot \Phi_{i+1,i}$$

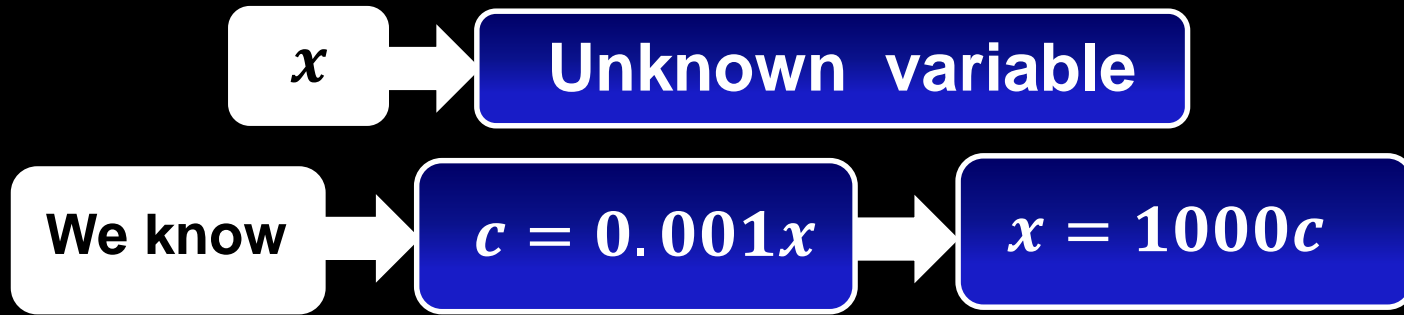
Ill-conditioned problem

Example 1: scalar form



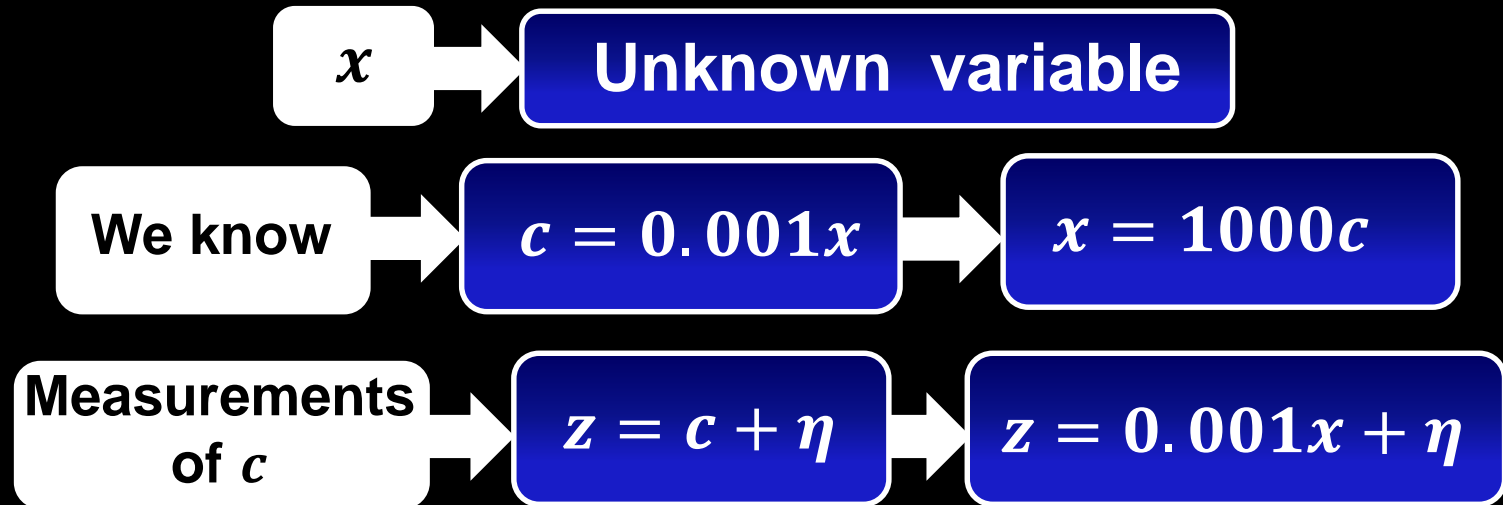
Ill-conditioned problem

Example 1: scalar form



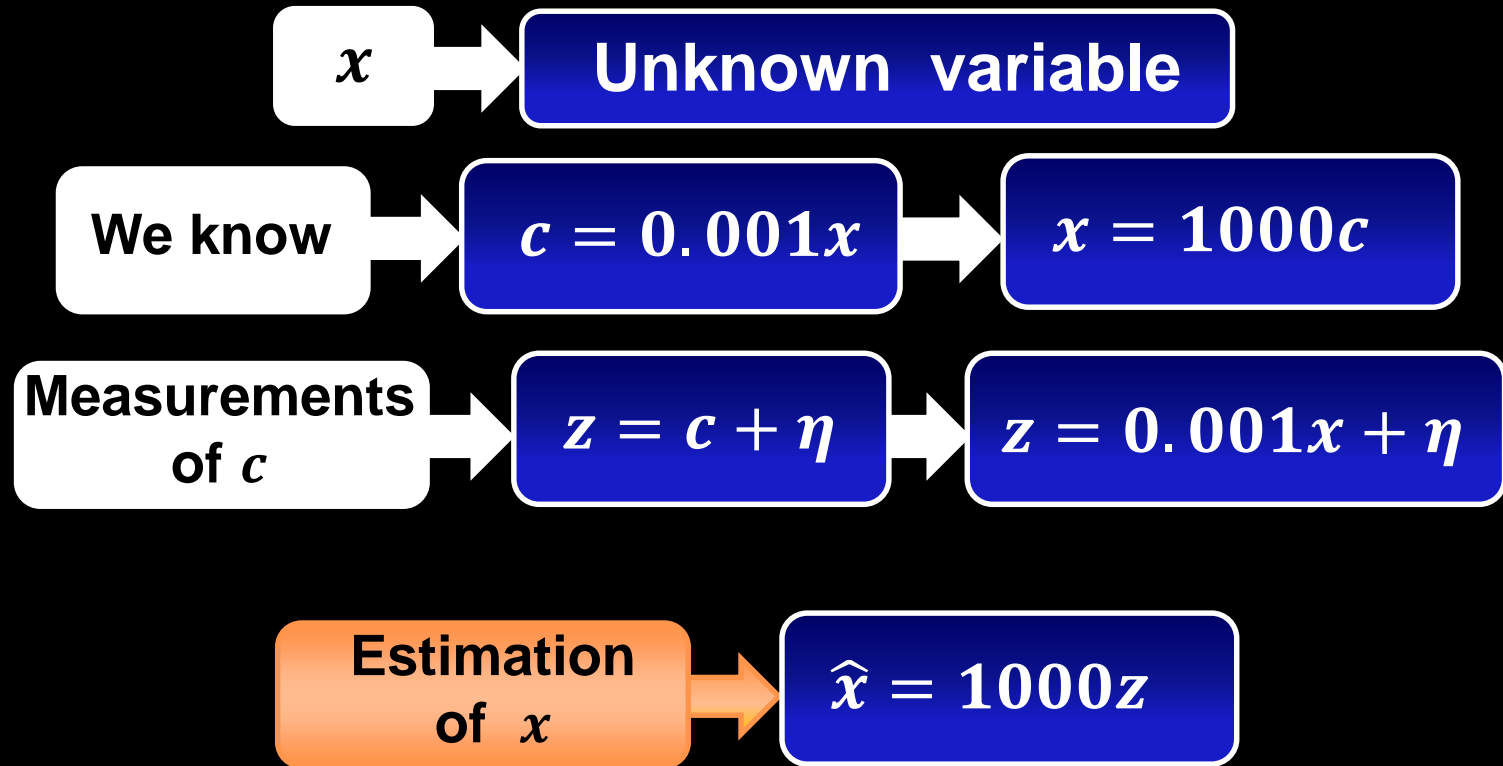
Ill-conditioned problem

Example 1: scalar form



Ill-conditioned problem

Example 1: scalar form



Ill-conditioned problem

Example 1: scalar form

x → Unknown variable

We know → $c = 0.001x$ → $x = 1000c$

Measurements of c → $z = c + \eta$ → $z = 0.001x + \eta$

Estimation of x → $\hat{x} = 1000z$

Estimation error → $x - \hat{x} = 1000(z - c)$ → $x - \hat{x} = 1000\eta$

Ill-conditioned problem

Example 1: scalar form

x → Unknown variable

We know → $c = 0.001x$ → $x = 1000c$

Measurements of c → $z = c + \eta$ → $z = 0.001x + \eta$

Estimation of x → $\hat{x} = 1000z$

Estimation error → $x - \hat{x} = 1000(z - c)$ → $x - \hat{x} = 1000\eta$

Estimation error is very high

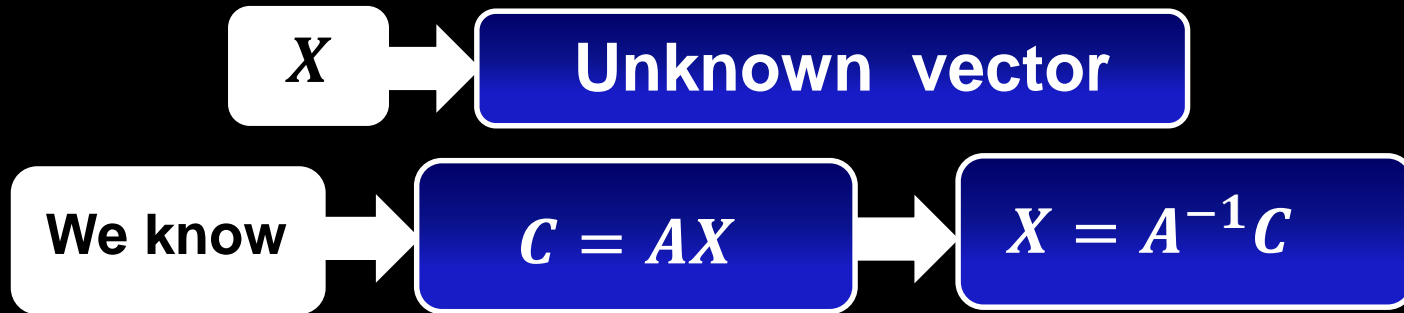
Ill-conditioned problem

Example 2: matrix form



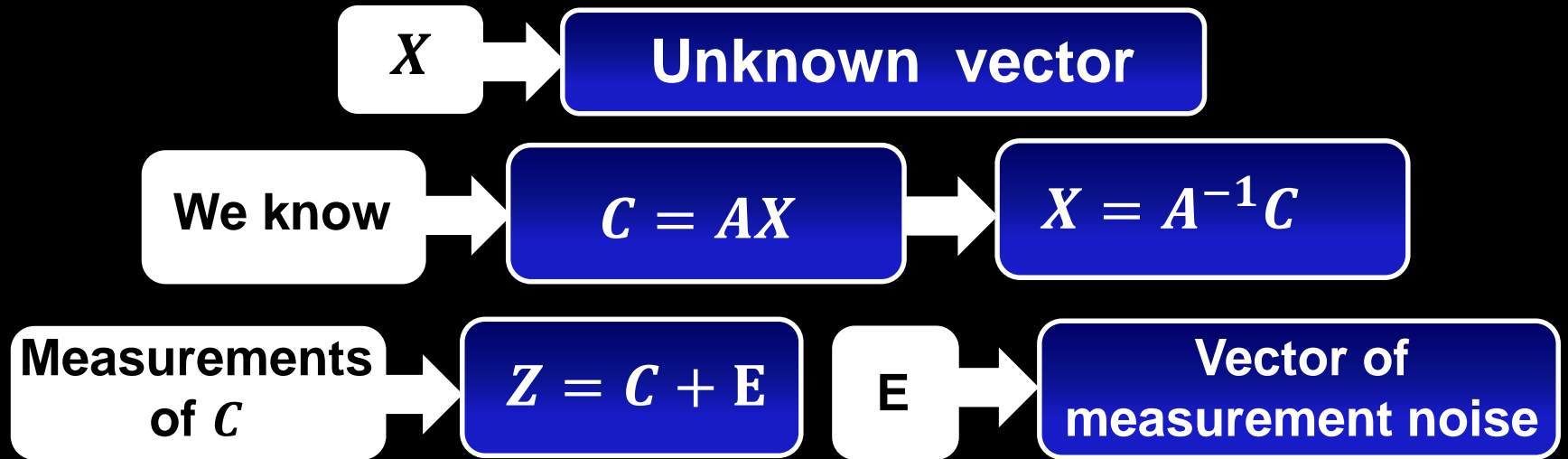
Ill-conditioned problem

Example 2: matrix form



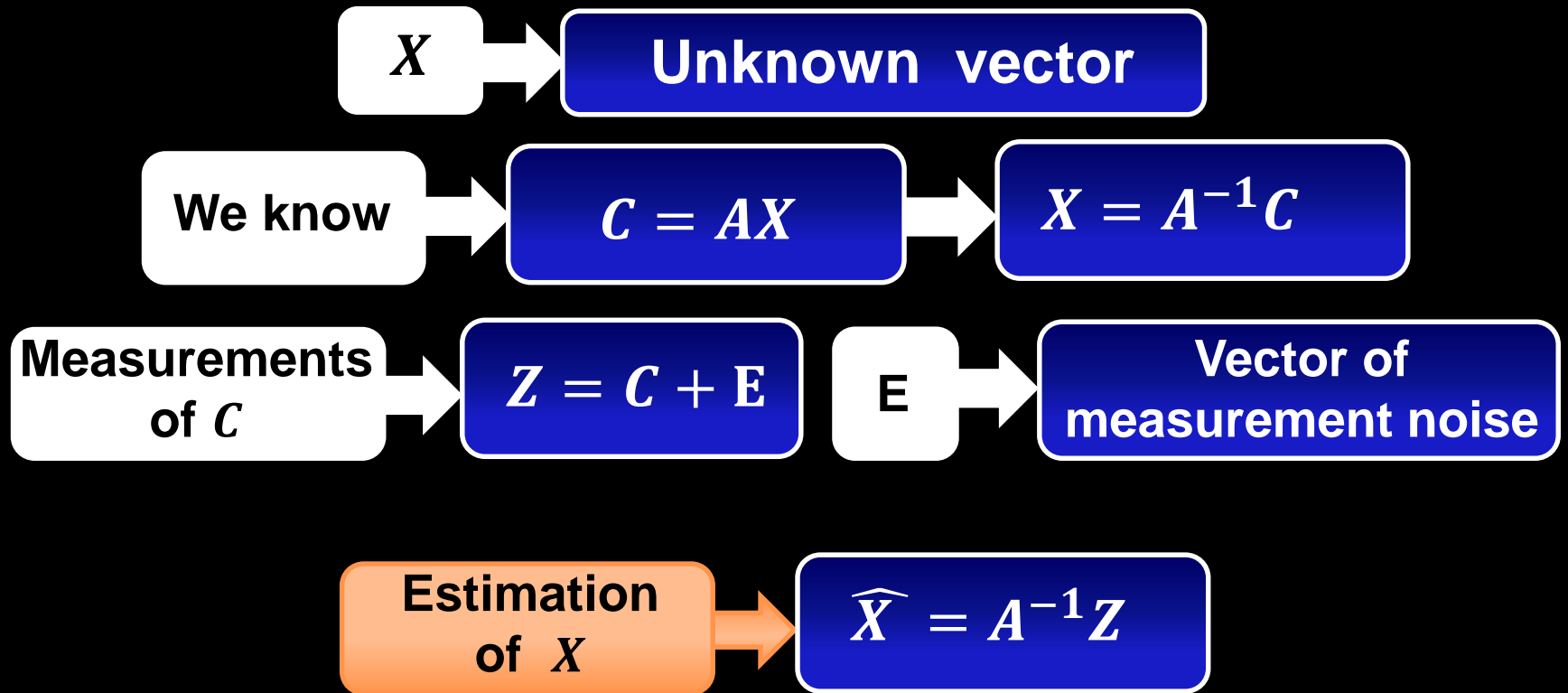
Ill-conditioned problem

Example 2: matrix form



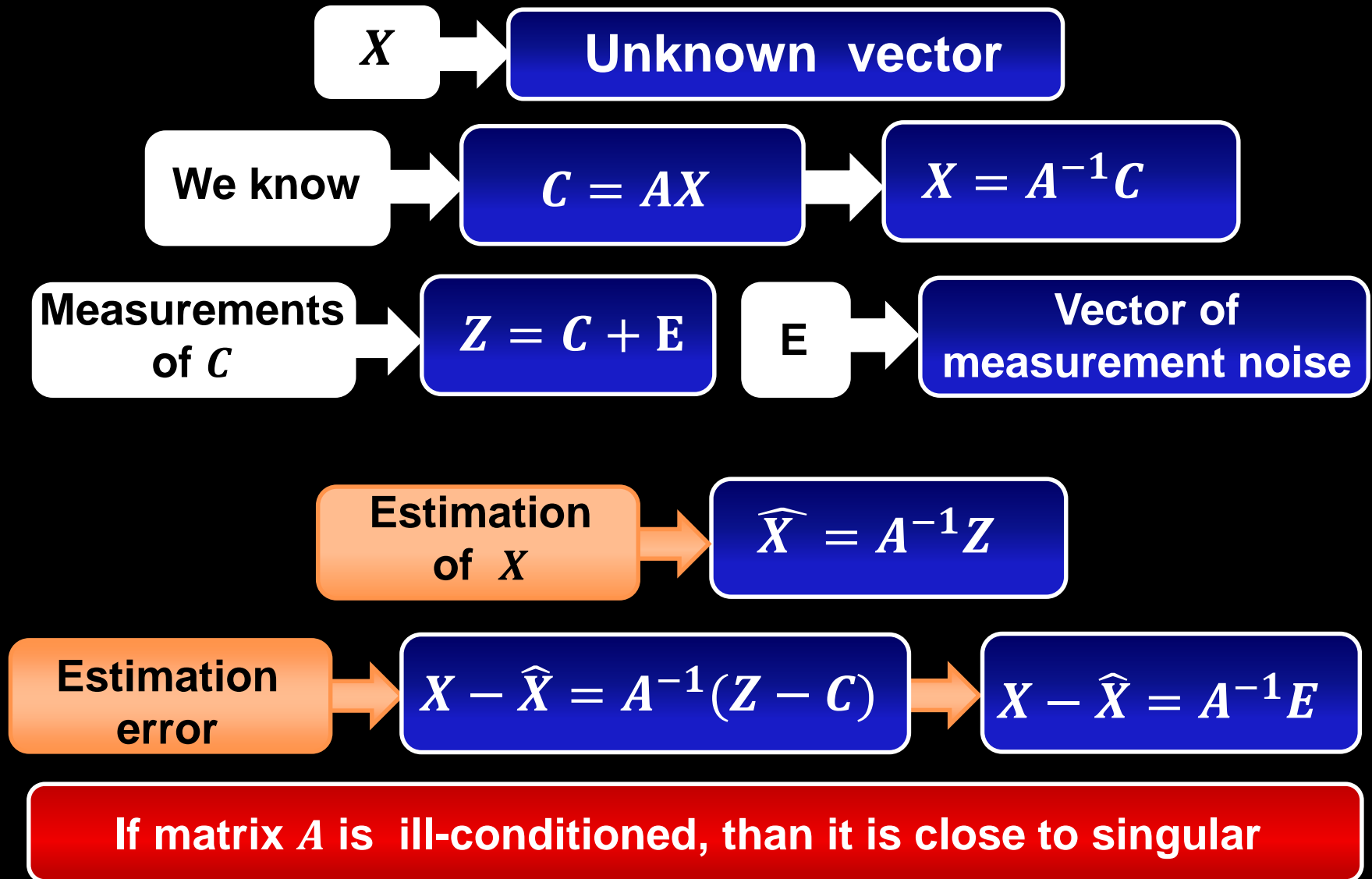
Ill-conditioned problem

Example 2: matrix form



Ill-conditioned problem

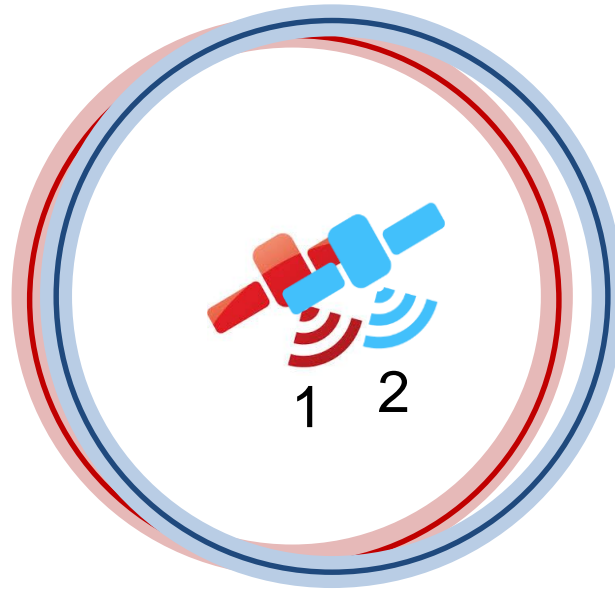
Example 2: matrix form



III-conditioned problem

1

**Validity
of applying
a technique**



-?



Man-made satellite



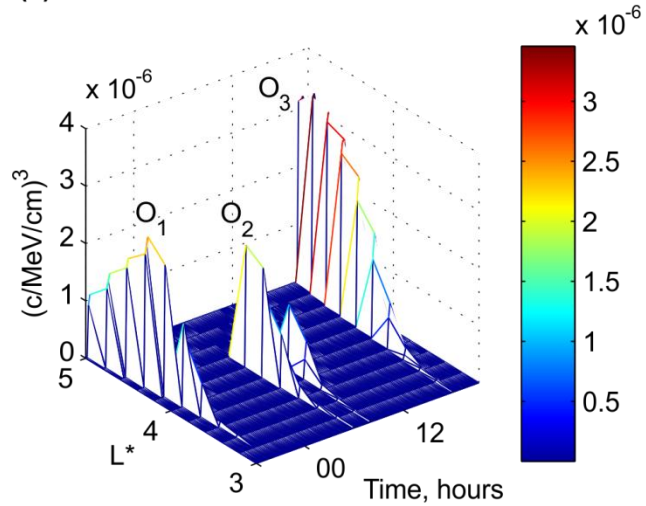
Navigation satellite

II-conditioned problem

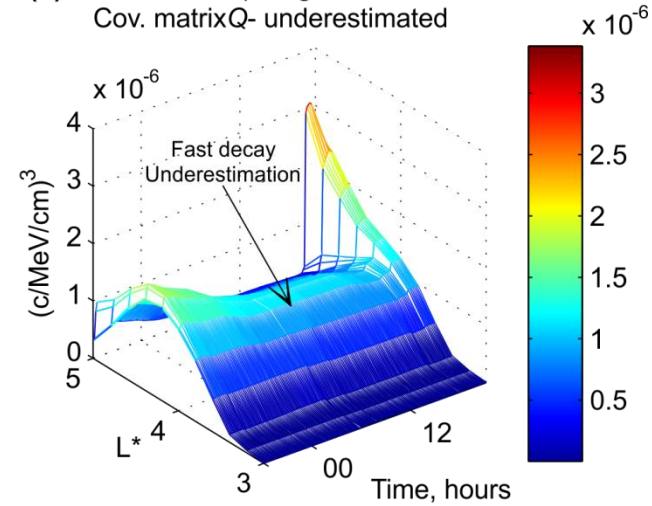
Satellite position is undefined!

Kalman filter needs noise statistics identification

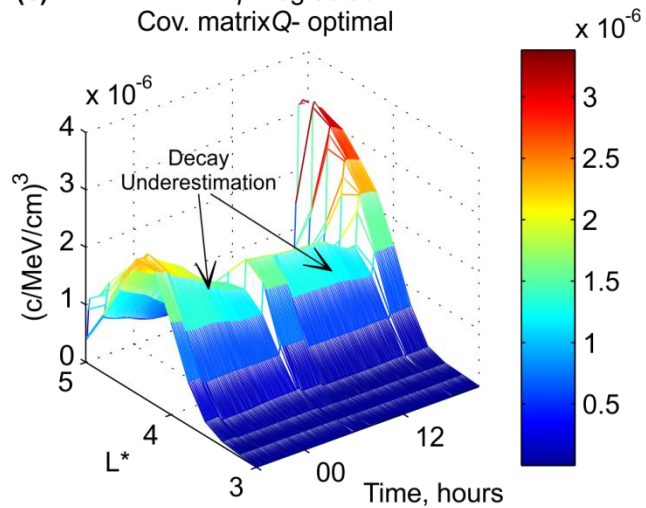
(a) The PSD observations



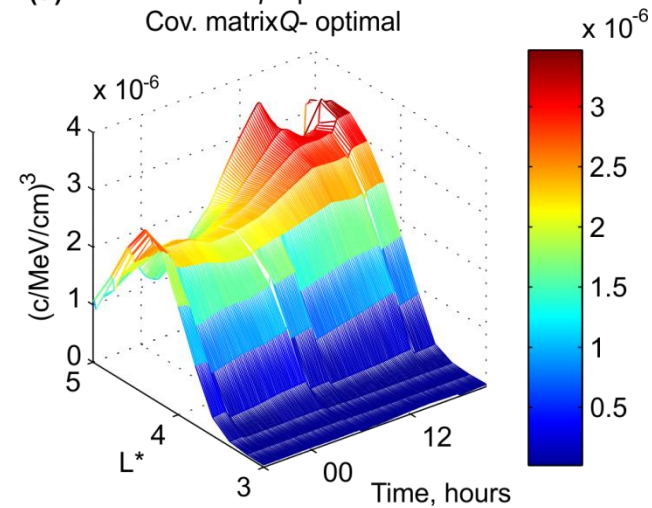
(b) Bias q -neglected
Cov. matrix Q -underestimated



(c) Bias q -neglected
Cov. matrix Q -optimal



(d) Bias q -optimal
Cov. matrix Q -optimal



Smoothing with fixed interval

Smoothing is performed in backward in time

$$X_{i,N} = X_{i,i} + A_i(X_{i+1,N} - \Phi_{i+1,i}X_{i,i})$$

$$i = N - 1, N - 2, \dots 1$$

Coefficient $A_i = P_{i,i}\Phi_{i+1,i}^T P_{i+1,i}^{-1}$

Smoothing error covariance matrix

$$P_{i,N} = P_{i,i} + A_i(P_{i+1,N} - P_{i+1,i})A_i^T$$

$X_{i,i}$ - filtered estimate, $X_{N,N}$ - initial estimate

$P_{i,i}$ - filtration error covariance matrix

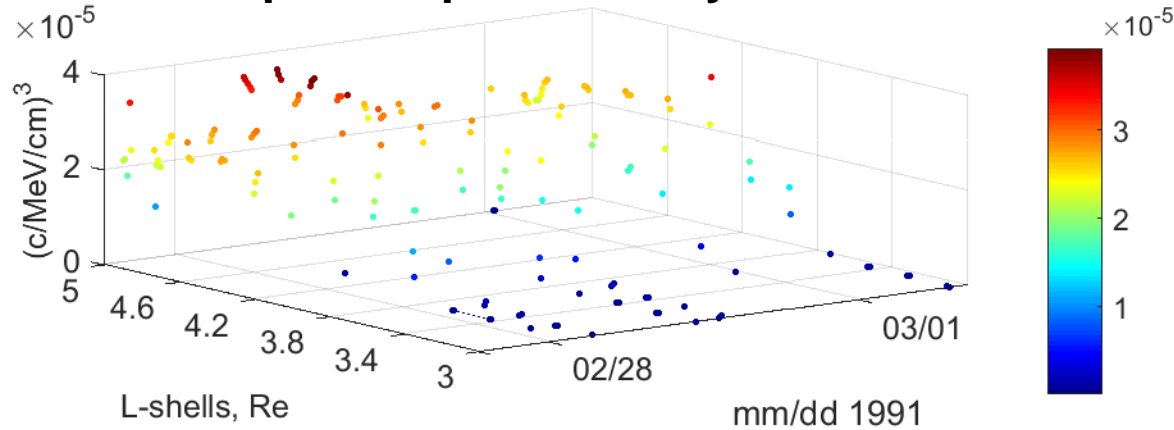
$P_{i+1,i}$ - prediction error covariance matrix

Smoothing takes into account both current and future measurements and therefore provides improved estimation

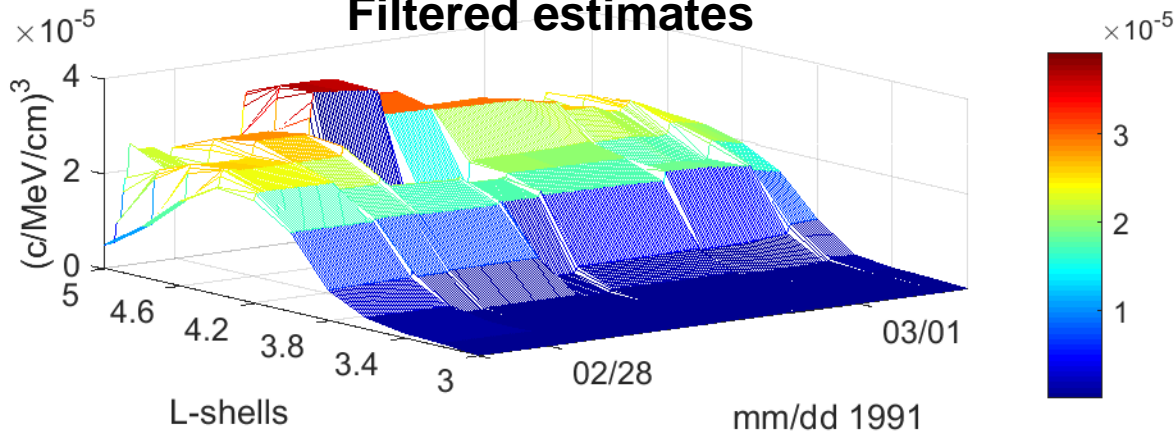
Reconstructing the dynamics of relativistic electrons in Earth's radiation belts

Smoothing takes into account both current and future measurements and therefore provides improved estimation

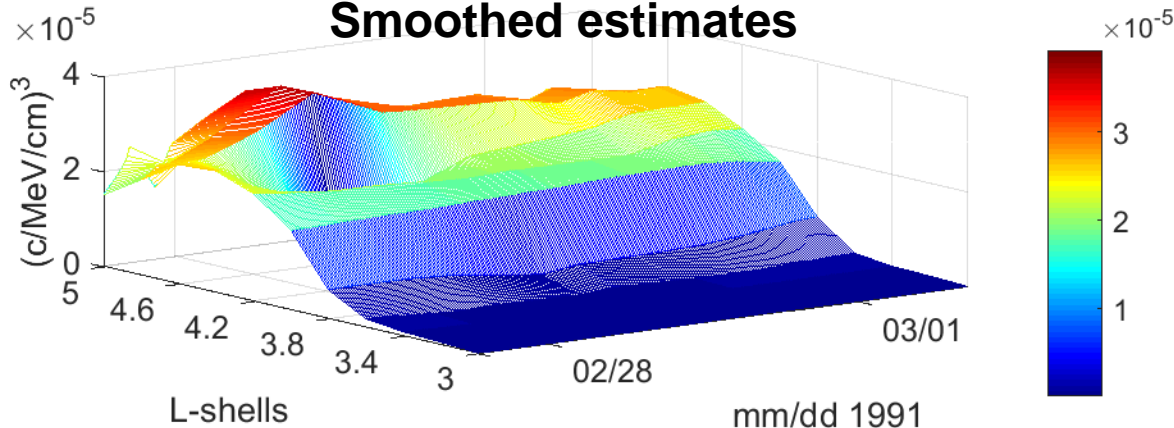
Electron phase space density observations



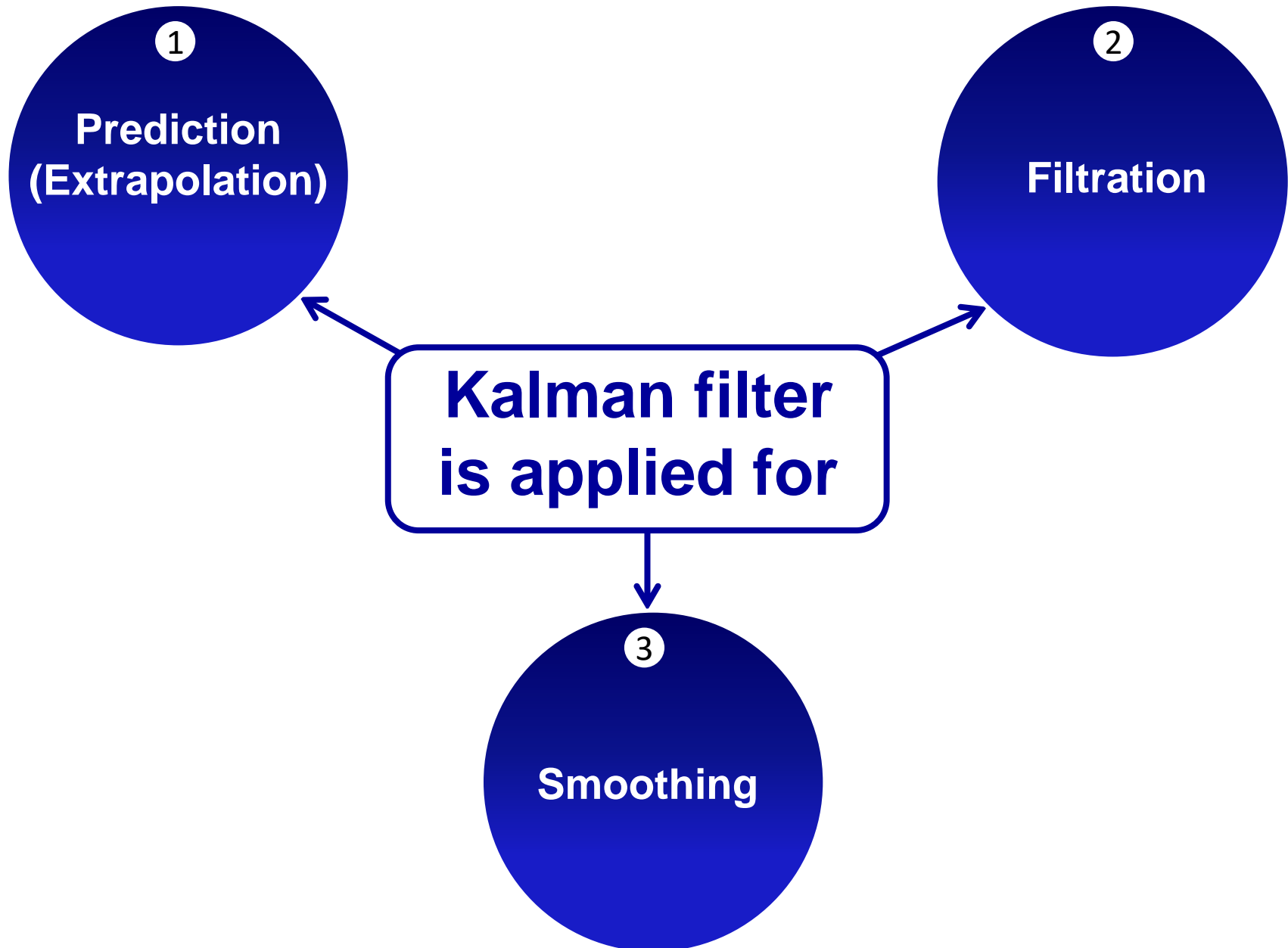
Filtered estimates



Smoothed estimates



Podladchikova et al. (2014), Noise statistics identification for Kalman filtering of the electron radiation belt observations:
2. Filtration and smoothing,
J. Geophys. Res. Space Physics, 119



Equivalence of exponential smoothing and stationary Kalman filter

Random
walk model

$$\begin{aligned}x_i &= x_{i-1} + w_i \\z_i &= x_i + \eta_i\end{aligned}$$

This is state space model with following parameters

$$X_i = |x_i|$$

State
vector

$$\Phi = 1$$

Transition
matrix

$$H = 1$$

Observation
matrix

Stationary
Kalman filter

$$x_{i,i} = x_{i-1,i-1} + K(z_i - x_{i-1,i-1})$$

Filter gain K
becomes a constant

Exponential
smoothing

$$x_i = x_{i-1} + \alpha(z_i - x_{i-1})$$

Optimal α

$$\alpha = K$$

Conclusions

**Kalman filter is effective
tool for estimation and
forecasting**

**However it requires
good hands for tuning**