Chapter #5 → GRAY - BOX sy. identification

→ #1: K.F.

→ #2: S.E.M.

G.B sy. ID. USING K.F.

KF. is NOT a sy. ID. method - it is a VARIABLE-ESTIMATION Approach (so. sensor/observer)

However, we can use it ALSO for G.B. sy. ID.

("SIDE BENEFIT" of KF.)

problem DEFINITION:

we have a model, typically BUILT as a W.B. model using first principles:

$$
\mathcal{S}: \begin{cases} x(t+1) = f\left(x(t), u(t); \vartheta\right) + v_1(t) & \text{- model noise} \\ y(t) = h\left(x(t); \vartheta\right) + v_2(t) & \text{- output/sensor noise} \end{cases}
$$

f and h are linear or NonLINEAR functions depending on some UN-KNOWN parameter ϑ (with a physical meaning ex: mass, RESISTANCE, friction coefficient,) --> problem --> estimate $\hat{\vartheta}$

K.F. solves this problem by transforming the UN-KNOWN parameters in EXTENDED STATES → K.F. makes the simultaneous estimation of $\hat{x}(t|t)$ (classic K.F. problem)

$\underline{\hat{\vartheta}(t)}$ (param. ident. problem)

↓ "<u>TRICK</u>": <u>STATE EXTENSION</u>:

$$S: \begin{cases} x(t+1) = f\left(x(t); u(t); \vartheta\right) + v_1(t) \\ \boxed{\vartheta(t+1) = \vartheta(t) + v_\vartheta(t)} \\ y(t) = h\left(x(t); \vartheta\right) + v_2(t) \end{cases}$$

the UN-KNOWN <u>PARAMETERS</u> are TRANSFORMED in UN-KNOWN <u>VARIABLES</u>

NEW "EXTENDED" STATE VECTOR is: $x_E = \begin{bmatrix} x(t) \\ \vartheta(t) \end{bmatrix}$

NEW EQUATION WE "CREATE"

> NOTICE: this equation is NOT of AN ASY. STABLE SYSTEM but of a SIMPLY-STABLE SYSTEM → No problem
KF can deal with NON-ASY STABLE SYSTEMS

- it is a "fictitious" equation ("TRICK")
(NOT a "physical" equation)

$$\vartheta(t+1) = \vartheta(t) + v_\vartheta(t)$$

↓ CORE DYNAMICS: $\vartheta(t+1) = \vartheta(t)$

this is the equation of something which is CONSTANT ⇒ this is exactly the NATURE of $\vartheta(t)$ which is INDEED a CONSTANT VECTOR of parameters

→ we NEED this "fictitious" noise in order to FORCE KF. To find (look for) the right VALUE of ϑ
(if NO NOISE in this equat. KF probably could stay fixed on the initial condition)
we tell KF → DO NOT RELY on i.c. !!!

Design choice is the choice of the covariance matrix of $v_{12}(t)$

$v_{12}(t) \triangleq WN(0, V_{12})$ vector of W.N.

Empirical assumption: $v_1 \perp v_{12}$; $v_2 \perp \cancel{v}_{12}$

(there is no reason for $v_{12}(t)$ to be correlated with v_1 and v_2)

$$V_{12} = \begin{bmatrix} \lambda_{1 2}^2 & & \varnothing \\ & \lambda_{2 2}^2 & \\ \varnothing & & \ddots & \\ & & & \lambda_{n_{12}}^2 \end{bmatrix}$$

← usually it is assumed that

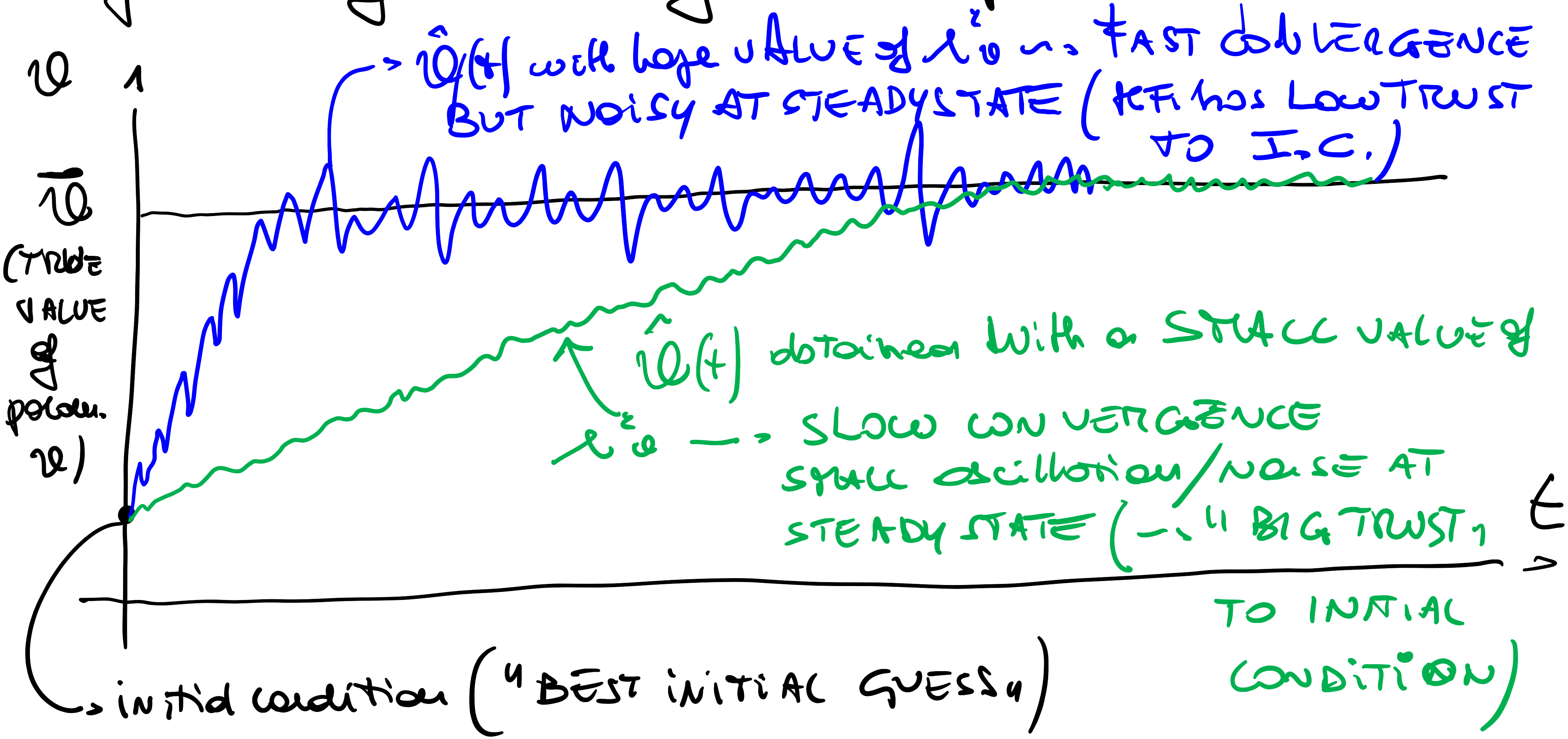
$\lambda_{1 2}^2 = \lambda_{2 2}^2 = \dots = \lambda_{n_{12}}^2 = \lambda_{12}^2$

⇒ we observe that $v_{12}(t)$ is a set of independent W.N. all with the same variance λ_{12}^2 → TUNING (DESIGN) PARAMETER

size $n_{12} \times n_{12}$

(n_{12} is the number of parameters in θ)

Influence of the choice of design param. λ_ϑ^2 →



→ λ_ϑ^2 is selected ACCORDING to the BEST compromise for your specific application

Notice: this "trick" can work in principle with any number of unknown parameters:

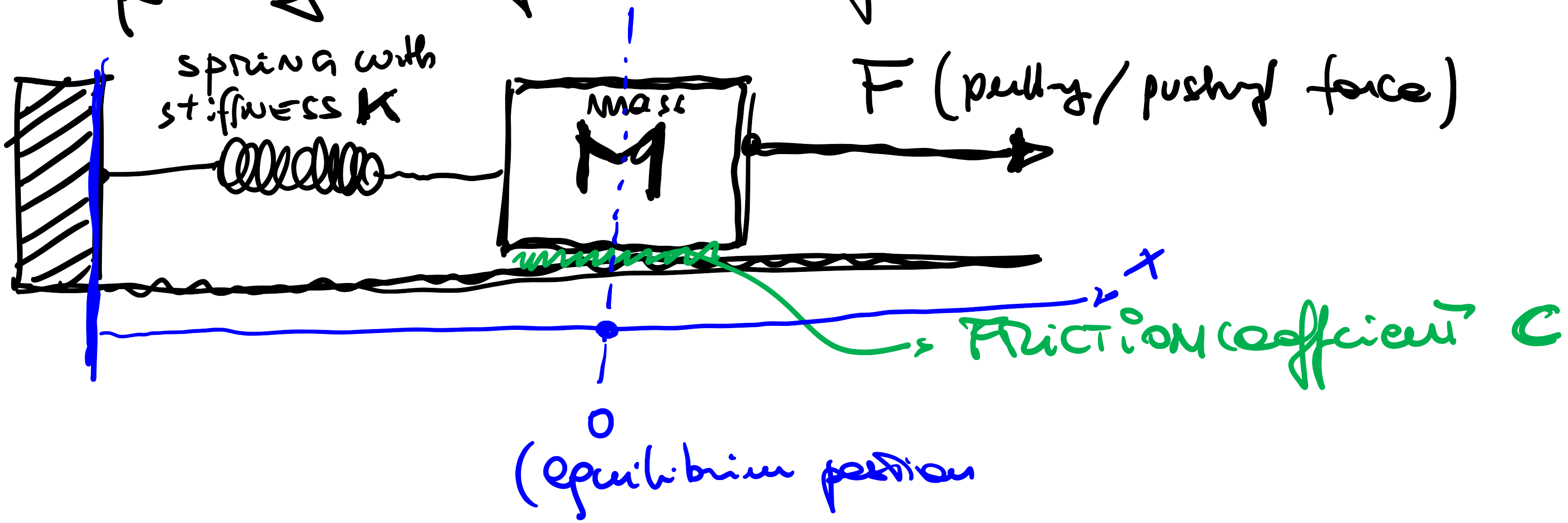
(eg. 3 sensors; 10 states; 20 parameters)

IN PRACTICE it works well only on a limited number of parameters:

↳ Ex: eg. "3 sensors; 5 states; 2 parameters"

(extreme way of using KF is → B.B. sy. identification)

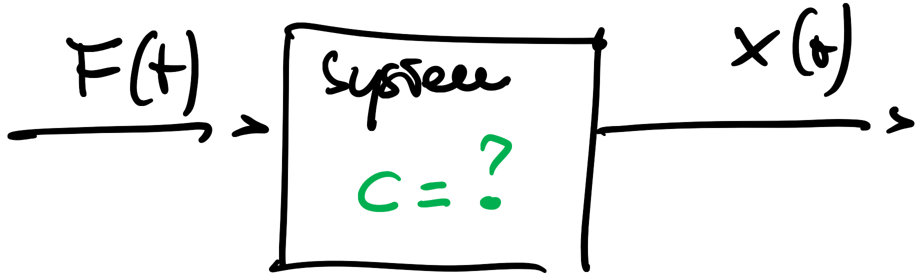
Example of KF for G.B. param. estimation:



Input variable: $F(t)$

output (measured) variable: position $x(t)$

parameters K and M are known (measured) but c is UN-KNOWN



USING K.F. we do NOT NEED a "TRAINING dataset"

problem → ESTIMATE c with KF

STEP #1 : model (W.B.) the system:

$$\boxed{\ddot{x}M = -Kx - c\dot{x} + F}$$

differential CONT. TIME linear eq.

↳ 2nd order eq --- we NEED 2 STATE VARIABLES

mech. syst ⟹ { position, speed }

$$x_1(t) = x(t)$$
$$x_2(t) = \dot{x}(t)$$

⇓

$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ M\dot{x}_2(t) = -Kx_1(t) - cx_2(t) + F(t) \\ y(t) = x_1(t) \end{cases} \qquad \begin{cases} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = -\frac{K}{M}x_1(t) - \frac{c}{M}x_2(t) + \frac{1}{M}F(t) \\ y(t) = x_1(t) \end{cases}$$

STEP 3 -> <u>DISCRETIZATION</u>:

EULERO forward: $\dot{x}(t) \simeq \dfrac{x(t+1) - x(t)}{\Delta}$

$\boxed{\Delta = \text{sample time}}$

$\downarrow$

$$\begin{cases} \dfrac{x_1(t+1) - x_1(t)}{\Delta} = x_2(t) \\ \dfrac{x_2(t+1) - x_2(t)}{\Delta} = -\dfrac{K}{M} x_1(t) - \dfrac{c}{M} x_2(t) + \dfrac{1}{M} F(t) \\ y(t) = x_1(t) \end{cases}$$

$\Downarrow$ STANDARD S.S. form: $\begin{cases} x(t+1) = Fx(t) + Gu(t) \\ y(t) = Hx(t) \end{cases}$

UN-KNOWN

$$\begin{cases} x_1(t+1) = x_1(t) + \Delta x_2(t) \\ x_2(t+1) = -\frac{K\Delta}{M} x_1(t) + \left(1 - \frac{c\Delta}{M}\right) x_2(t) + \frac{\Delta}{M} F(t) \\ y(t) = x_1(t) \end{cases}$$

LINEAR 2nd order discrete time system

STEP # 3 --> STATE EXTENSION:

$$\boxed{x_3(t) = c(t)}$$

$$
\begin{cases}
x_1(t+1) = x_1(t) + \Delta x_2(t) + v_{11}(t) \\
x_2(t+1) = -\frac{k\Delta}{m} x_1(t) + \left(1 - \frac{\Delta x_3(t)}{M}\right) x_2(t) + \frac{\Delta}{M} F(t) + v_{12}(t) \\
x_3(t+1) = x_3(t) + v_{13}(t) \\
y(t) = x_1(t) + v_2(t)
\end{cases}
$$

NON-LINEAR operator

$$
V_1 = \begin{bmatrix} \lambda_{11}^2 & 0 & 0 \\ 0 & \lambda_{12}^2 & 0 \\ 0 & 0 & \lambda_{13}^2 \end{bmatrix}
$$

$$
V_2 = \lambda^2
$$

READY for K.F. Application → we get at the same time $\hat{x}(t)$ and $\hat{c}(t)$

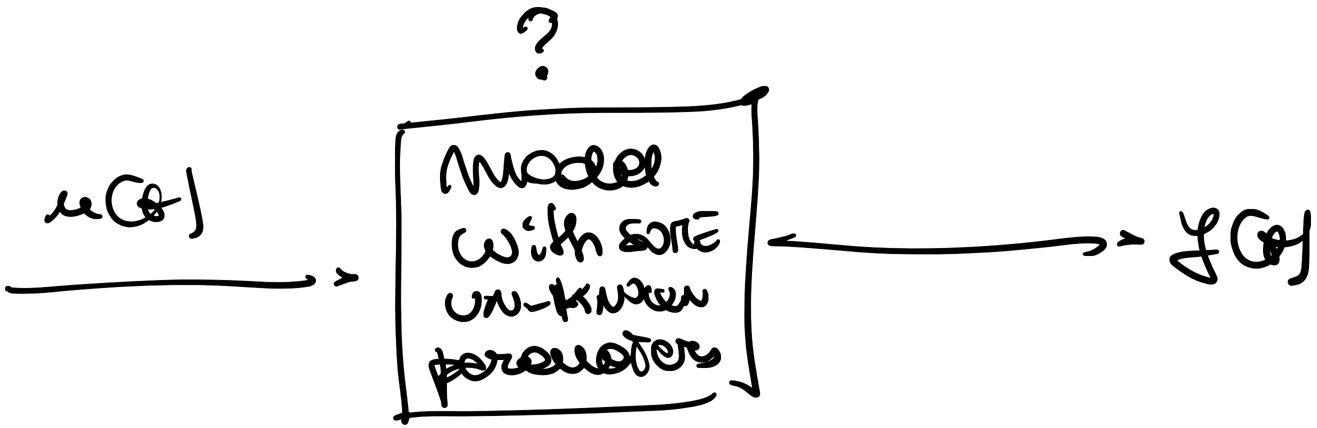
Notice that we NEED EKF → even if
original system was linear → state
extension → NON LINEAR!

Are there ALTERNATIVE ways to solve GRAY-BOX sy. ID. problems?

YES → commonly (intuitive) used method is **PARAMETRIC** identification approach based on

SIMULATION
ERROR
METHOD

SEM



1) Collect data from an experiment $\longrightarrow$

$$\{\tilde{u}(1), \tilde{u}(2) \dots \tilde{u}(N)\}$$

$$\{\tilde{y}(1), \tilde{y}(2) \dots \tilde{y}(N)\}$$

2) DEFINE model structure:

$$y(t) = \mathcal{M}\left(u(t); \bar{\vartheta}; \vartheta\right)$$

$\bar{\vartheta}$ $\to$ set of KNOWN parameters (mass, resistance, etc.)

ϑ $\to$ set of UN-KNOWN parameters (possibly with BOUNDS $\vartheta_{min} < \vartheta < \vartheta_{max}$)

$u(t)$: input signal

$\mathcal{M}$: mathematical model (linear or nonlinear) usually written from first principle equations

3) perf. index definition: SEM:

$$J_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} \Big(\underbrace{\tilde{y}(t)}_{\text{measured output (true output)}} - \underbrace{m\big(\overset{\text{measured input}}{\tilde{u}(t)}; \bar{\theta}; \theta\big)}_{\text{simulated output}} \Big)^2$$

sample variance of the simulation error

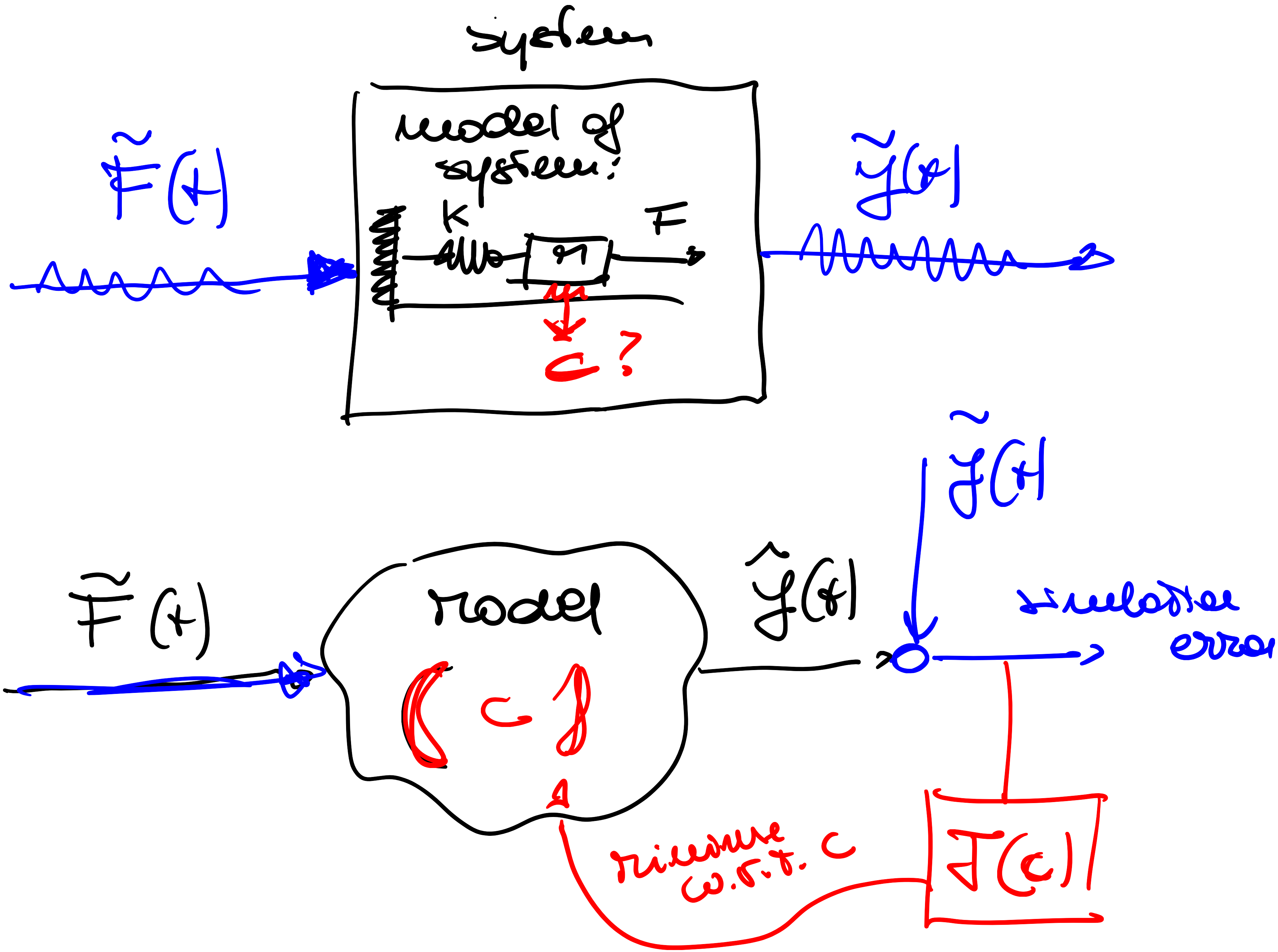
4) optimization:

$$\hat{\vartheta}_N = \arg\min_{\vartheta} \{ J_N(\vartheta) \}$$

Usually:
- NO ANALYTIC expression of $J_N(\vartheta)$ is AVAILABLE

Intuitive but comput. very demanding

- EACH computation of $J_N(\vartheta)$ requires an <u>entire simulation</u> of the <u>model</u> from $t=1$ to $t=N$
- Usually $J_N(\vartheta)$ is a non-quadratic and NON CONVEX function ⇒ ITERATIVE and randomized optim. methods must be used.


system
model of system:
K
M
F
c ?
$\tilde{F}(t)$
$\tilde{y}(t)$
model
$\{c\}$
$\hat{y}(t)$
$\tilde{y}(t)$
simulation error
$J(c)$
minimize w.r.t. c

$\boxed{\text{Question! can S.E.M. be applied also to B.B. methods ?}}$

Consider the following example:

we collect data: $\{\tilde{u}(1) \ldots \tilde{u}(N)\}$ $\{\tilde{y}(1) \ldots \tilde{y}(N)\}$

we want to estimate from data this I/O model:

$$\boxed{y(t) = \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}} u(t-1)} \qquad \theta = \begin{bmatrix} a_1 \\ a_2 \\ b_0 \\ b_1 \end{bmatrix}$$

↓ TIME DOMAIN:

$$y(t) = -a_1 y(t-1) - a_2 y(t-2) + b_0 u(t-1) + b_1 u(t-2)$$

Using P.E.M. $\rightarrow$

$$\hat{y}(t|t-1) = -a_1 \tilde{y}(t-1) - a_2 \tilde{y}(t-2) + b_0 \tilde{u}(t-1) + b_1 \tilde{u}(t-2)$$

MEASURED DATA

$\downarrow$

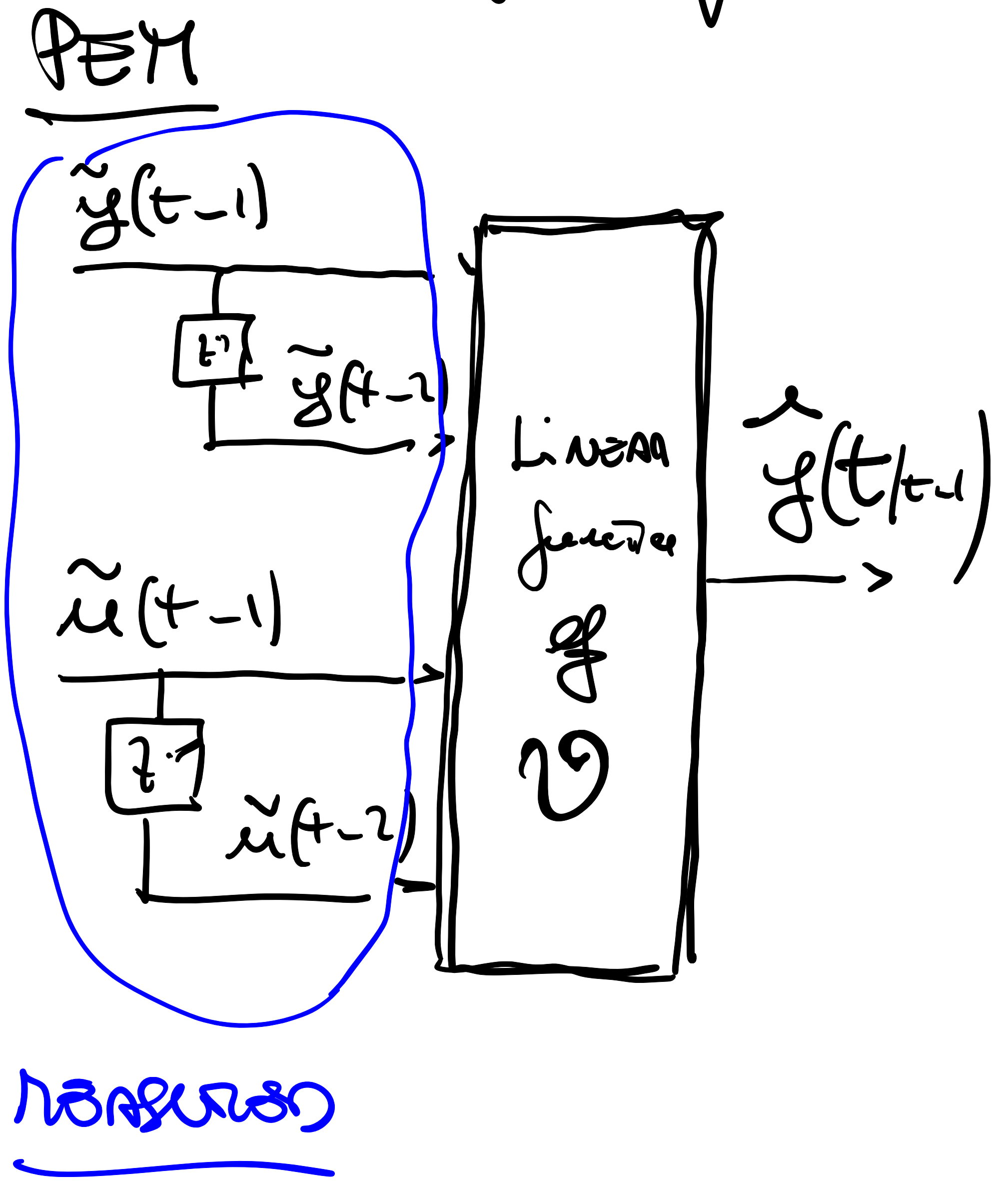
$$J_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} \left(\tilde{y}(t) - \hat{y}(t|t-1;\theta) \right)^2 =$$

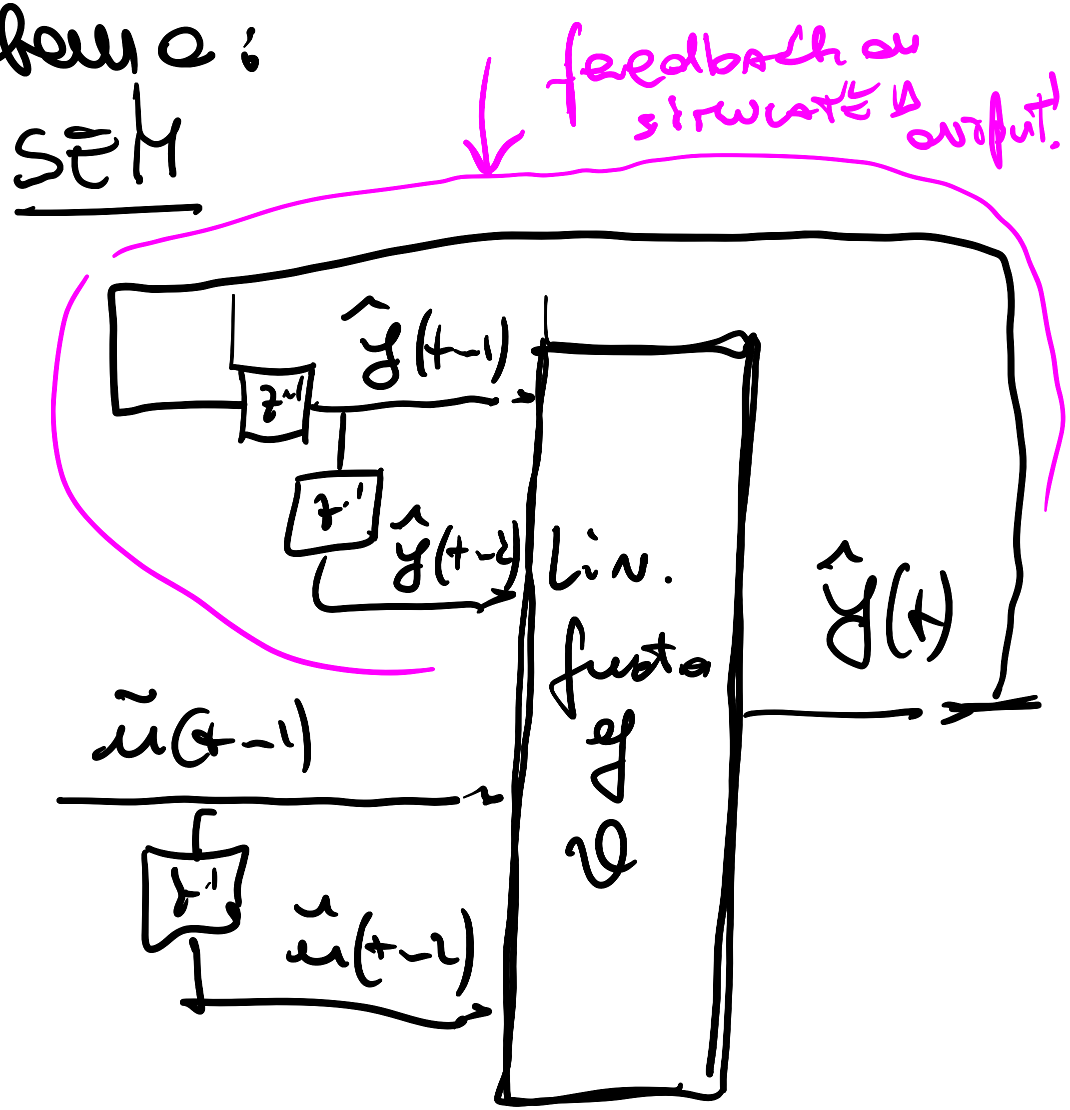
P.E.M.

$$\boxed{= \frac{1}{N} \sum_{t=1}^{N} \Big(\tilde{y}(t) + \underbrace{a_1 \tilde{y}(t-1) + a_2 \tilde{y}(t-2) - b_0 \tilde{u}(t-1) - b_1 \tilde{u}(t-1)}_{\text{LINEAR W.R.T } \theta} \Big)^2}$$

↳ QUADRATIC function of θ! $\rightarrow$ minimization is very simple

↓ corresponding Block schema:





SEM:

$$\hat{y}(t) = -a_1 \hat{y}(t-1) - a_2 \hat{y}(t-2) + b_0 \tilde{u}(t-1) + b_1 \tilde{u}(t-2)$$

old values of the simulated output

MEASURED

$$J_N(\vartheta) = \frac{1}{N} \sum_{t=1}^{N} \left(\tilde{y}(t) - \hat{y}(t;\vartheta) \right)^2 =$$

SEM

$$= \boxed{\frac{1}{N} \sum_{t=1}^{N} \left(\tilde{y}(t) + a_1 \hat{y}(t-1) + a_2 \hat{y}(t-2) - b_0 \tilde{u}(t-1) - b_1 \tilde{u}(t-2) \right)^2}$$

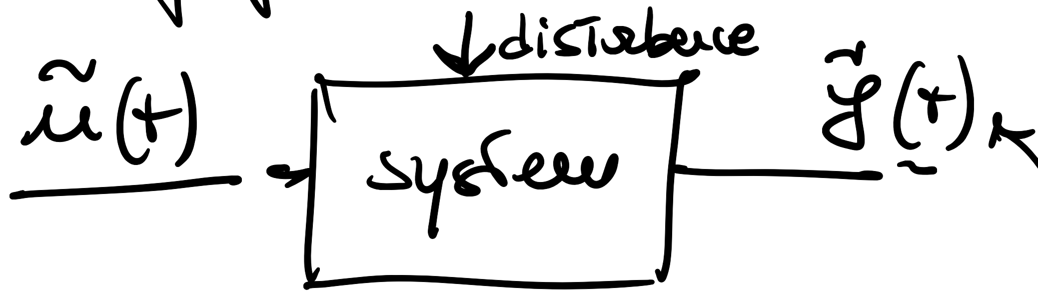
is highly NONLINEAR w.r.t. ϑ

=> NON QUADRATIC / NON CONVEX

PEM approach looks much BETTER ! (?)

$\boxed{\text{BUT}}$:

1) DO NOT forget the noise!



$\tilde{y}(t)$ also contains some disturbance!

PEM is much LESS ROBUST w.r.t. to NOISE $\Longrightarrow$

we **MUST** include a model of the noise in the estimated model $\longrightarrow$ $\boxed{\text{ARMA}}$X models

noise model

if ARX: $\quad e \sim WN$

$$y(t) = \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}} u(t-1) + \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2}} e(t)$$

$$\downarrow$$

$$\hat{y}(t|t-1) = \underbrace{b_0 u(t-1) + b_1 u(t-2) - a_1 y(t-1) - a_2 y(t-2)}_{\text{LINEAR in the param. VECTOR.}}$$

if ARMAX $\quad$ NOT 1 BUT $\longrightarrow$ $1 + c_1 z^{-1} + \dots c_m z^{-n}$

$$\Downarrow$$

$J_N(\theta)$ is HIGHLY NONLINEAR $\longrightarrow$ SAME complexity of S.E.M.

2) VERY SENSITIVE TO sampling time choice:

Remember that when we write at discrete time $y(t)$ → we mean $y(t \cdot \Delta)$

$\uparrow$

"t times the sampling time Δ"

$$\hat{y}(t|t-1) = -a_1 \tilde{y}(t-1) - a_2 \tilde{y}(t-2) + b_0 \tilde{u}(t-1) + b_1 \tilde{u}(t-2)$$

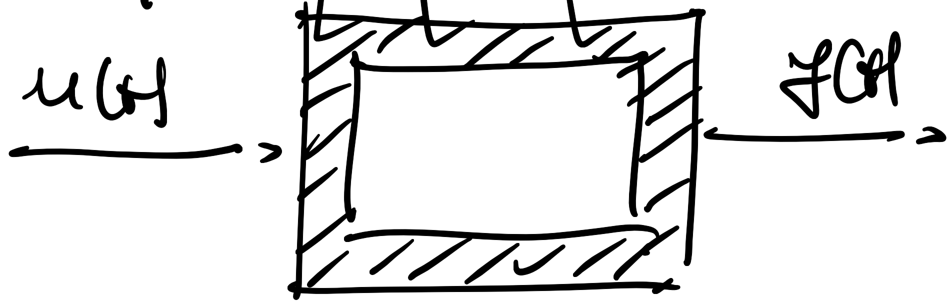
if Δ is very small ⟹ the difference between $\tilde{y}(t)$ and $\tilde{y}(t-1)$ becomes very small → $\emptyset$

$\Rightarrow$ effect that the P.E.M. optimization tends to provide this "TRIVIAL" solution:

$$\left.\begin{array}{l} a_1 \longrightarrow -1 \\ a_2 \longrightarrow \emptyset \\ b_0 \longrightarrow \emptyset \\ b_1 \longrightarrow \emptyset \end{array}\right\} \Rightarrow \boxed{\tilde{y}(t) \approx \tilde{y}(t-1)}$$

↳ wrong model due to the fact that the recursive part of the model is using past measures of the output instead of past values of the model outputs!!

Summary of system identi methods for I/O systems



- Collect a DATA SET for TRAINING (if NEEDED)
- choose a model DOMAIN
  - LINEAR STATIC
  - NON-linear STATIC
  - Linear dynamic
  - Nonlinear dynamic

  GRAY-BOX or B.BOX
- ESTIMATION method
  - CONSTRUCTIVE (4SID)
  - Parametric optimization
    - PEM
    - SEM
  - filtering (STATE EXTENSION of K.F)

Better Black Box
for syl Ident.
or sw. sensy

OR white Box ?

No simple answer / depends on goals and Type of Applications

B.B → VERY general
VERY FLEXIBLE
MAKE maximum USE of data
NO OR LITTLE NEED of DOMAIN KNOW-how

WB → VERY USEFUL when you are the system-designer (not only the control algo. designer !)
→ PROVIDE MORE INSIGHT in the system

Gray box can sometimes be obtained by Hybrid systems →

part is B.B.

part is W.B.