DATA DRIVEN CONTROL SYSTEM DESIGN

<u>Nominal design</u>: does not guarantee <u>any robustness</u> as for reference tracking

↓

satisfying requirements fo a range of possible situations

<u>Integrators</u>: achieves robustness as for reference tracking, but it is conservative as for transient (long assessment time)

$$\boxed{\frac{k \cdot z}{z-1}}$$

k = gain

Control specifications

- zero tracking error
- small enough transient

+

high level <u>uncertainty</u>

multitude of possible systems <u>simultaneously</u>

ADAPTIVE CONTROL

(NON-LINEAR)

we implement the structure of the nominal controller but <u>the parameters ar</u> <u>not fixed</u> (they are continuously updated by the information carried over by data)

ADAPTATION RULE

Least Squares estimation

$$x(t+1) = ax(t) + u($$

$$\downarrow$$

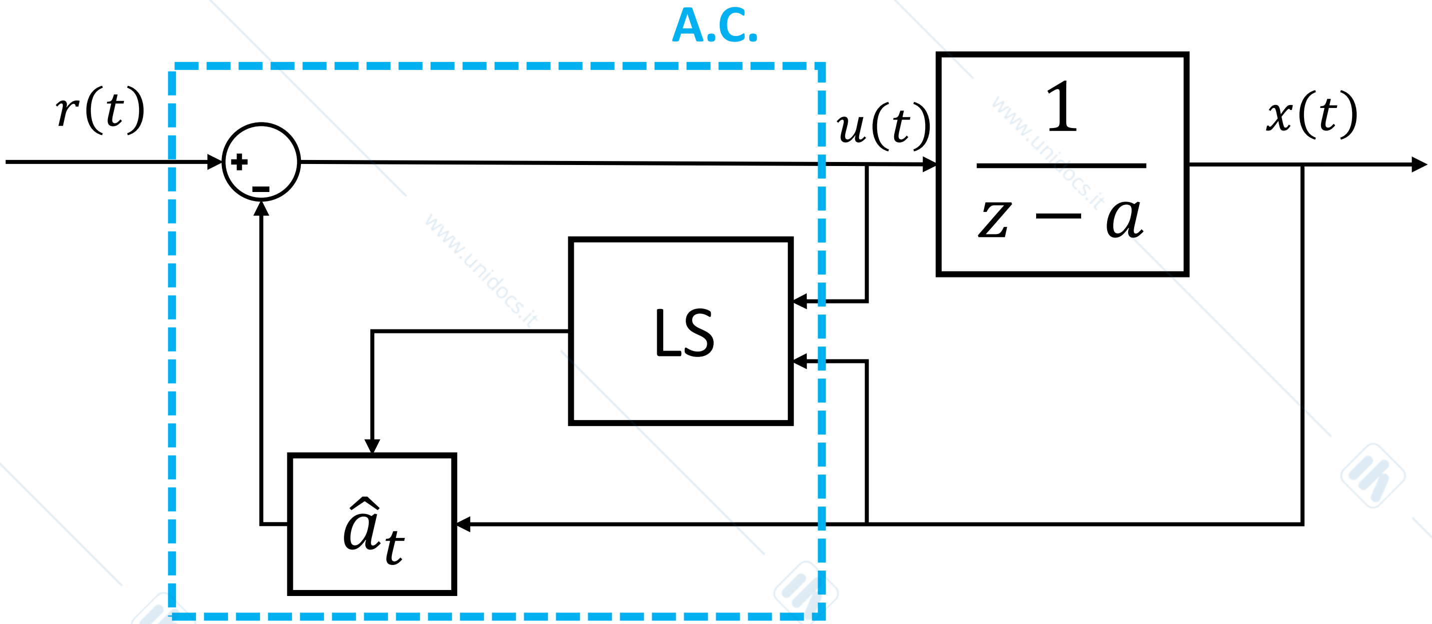
$$x(t+1) - u(t) = ax($$

$$\downarrow$$

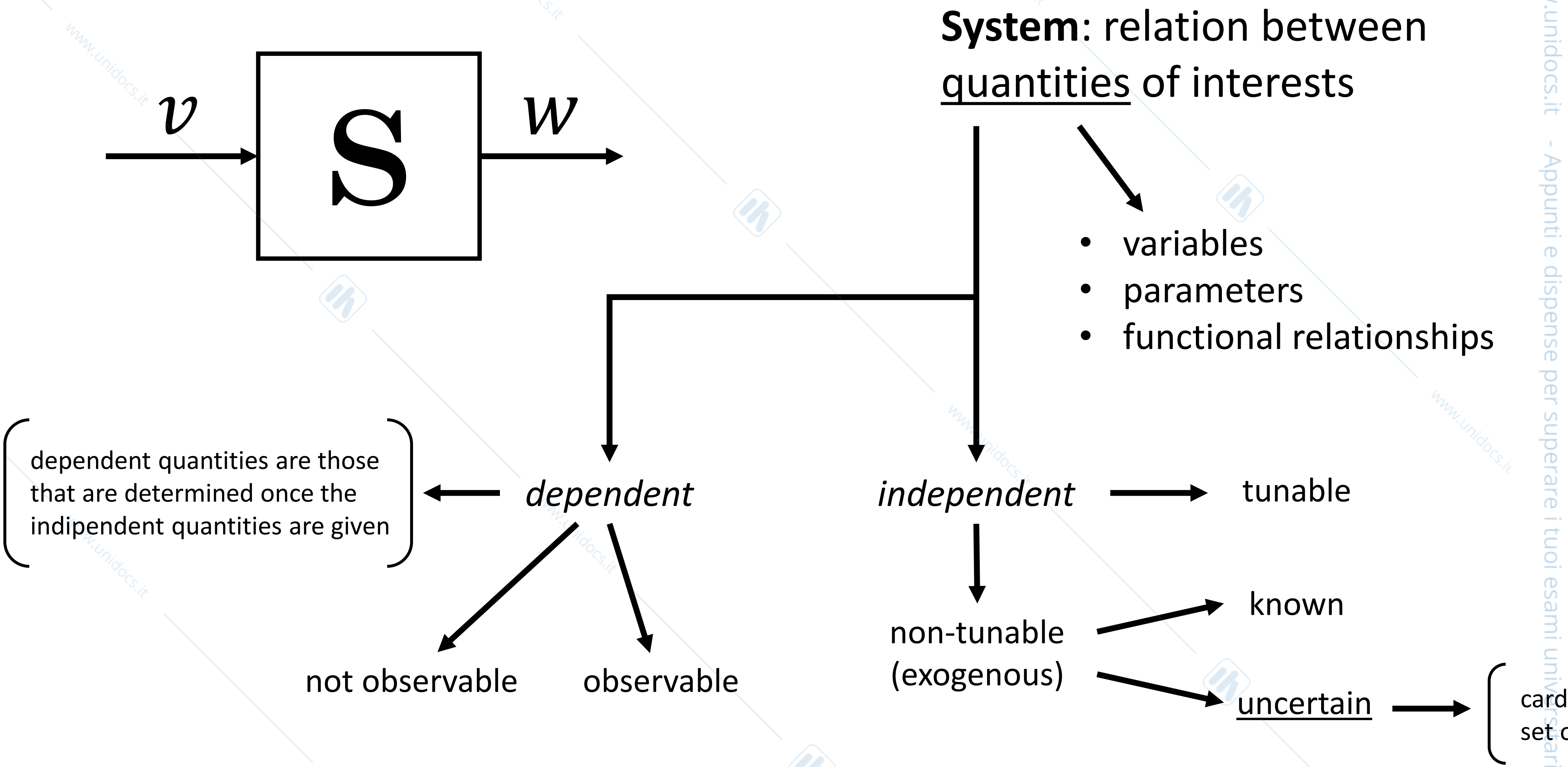
$$y(t+1) = a\varphi(t)$$

$$\hat{a}_{t+1} = \frac{\alpha\bar{a} + \sum_{i=0}^{t} \varphi(i) y(i+1)}{\alpha + \sum_{i=0}^{t} \varphi(i)\varphi(i)^T}$$

$$u(t) = -\hat{a}x(t) + r(t)$$



GENERAL PRINCIPLES



CONTROL PROBLEM

choose the value of <u>tunable</u> indipendent quantities based on the observable dependent quantities so that the behavior of the dependent quantities satisfies some given goal <u>irrespective of the values taken by the uncertain quantities</u> → ***robustness requirements***

CONTROLLER C → link between tunable and observable quantities

$\in$

CLASS $\mathcal{C}$

S → <u>realization</u> of a system for a given value of the uncertain quantities

$\in$

$\mathcal{S}$ → class of «systems» obtained for the various possible values of uncertain quant

$J(C,S) \longrightarrow$ performance index

control goal $\longrightarrow$ $J(C,S) \leq K$

- poles
- tracking error at $t \to \infty$

control requirements:
- stability
- settling time
- zero tracking error

goal $\longrightarrow$ find $C \in \boldsymbol{\mathcal{C}}$: $J(C,S) \leq K$, $\quad \forall S \in \boldsymbol{\mathcal{S}}$

(1) C is big enough (no need of adaptive control)

$$\exists C \in \boldsymbol{\mathcal{C}} : J(C,S) \leq K, \quad \forall S \in \boldsymbol{\mathcal{S}}$$

remarks:
- class of uncertain syste
- control requirements a strict (classical control robust control, ecc)
- unique controller

$$(2) \quad \forall C \in \mathcal{C} \ \exists S \in \mathcal{S} : \ J(C,S) > K$$

control problem cannot be solved within the chosen controller class → <u>change the controller cl</u>

$$\underbrace{\mathcal{C}}_{\text{linear controllers}} \longrightarrow \underbrace{\mathcal{C}'}_{\text{non linear}} \longrightarrow$$

enlarged class of controllers that allows to solve the control problem (must be not too big otherwise we are no more able to operate with it)

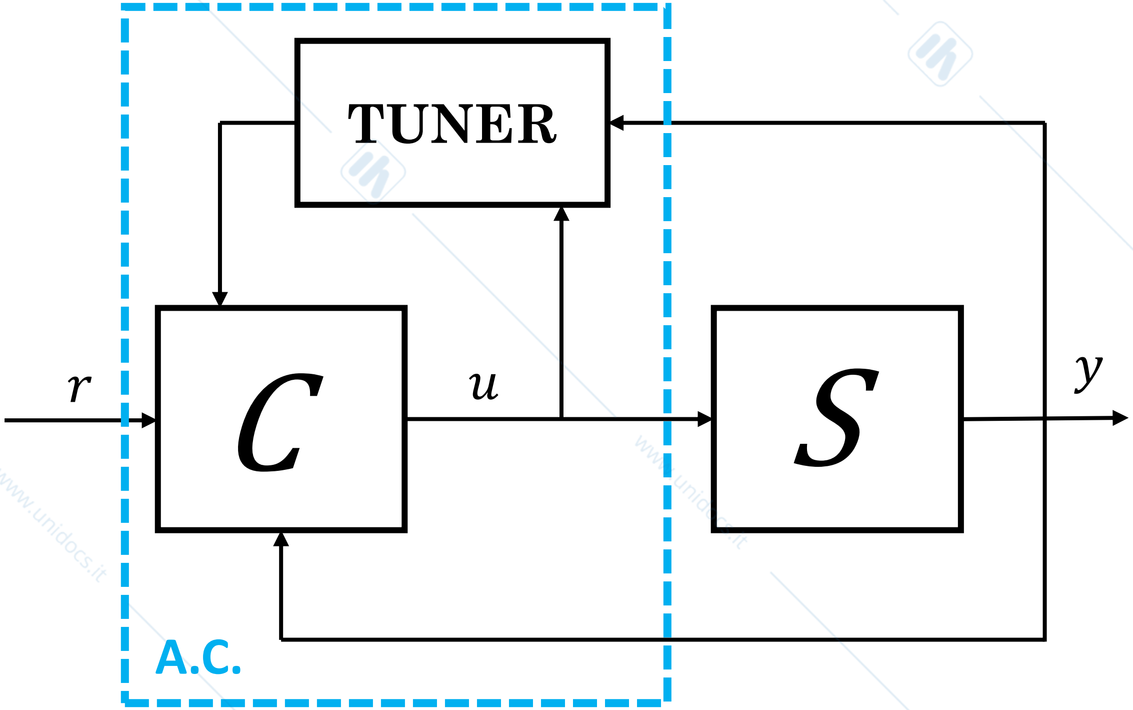
↓

adaptive control is nothing but a means to <u>construct a suitable</u> $\mathcal{C}'$ from $\mathcal{C}$

$$(2.1) \quad \forall S \in \mathcal{S} \ \exists C \in \mathcal{C} : \ J(C,S) \leq K$$

if we fix uncertainty, then we are capable to solve <u>one control problem</u> (different S lead to different C in this condition)

Adaptation rule: we implement the structure of a controller in C, but it is <u>not kept fixed</u>, it is <u>tuned</u> through time based on the <u>a-posteriori information</u> obtained from input and output S (the response of S, value taken by the dependent observable quantities given the value of indipendent tunable quantities)



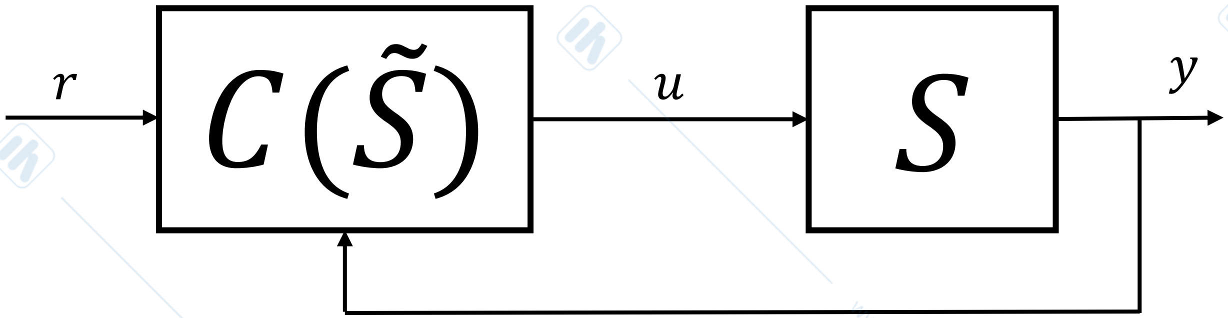
<u>UNCERTAINTY</u> calls for **ADAPTATIO**
no matter if the system is t-invarian
or t-variant (but often A.C. is the on
possibility to solve the robust contr
problem of t-variant systems)

mechanism through which S is updated i
adaptation rule (to counteract the uncert

enlarged class of controllers beacuse of the **<u>tuner</u>**

C generates from $\longrightarrow$

- a-posteriori info (observations of observable quanti
- a-priori info on S ($\tilde{S}$ any knowledge about quantitie the relation among quantities in S)



- a-posteriori info only $\Rightarrow$ <u>controller is t-invariant</u>
- a-priori info is what may justify the use of t-varying controllers

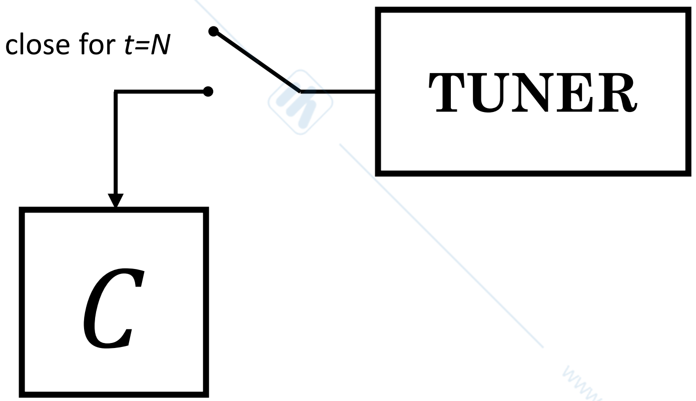
TAXONOMY OF ADAPTIVE CONTROLLERS	**INDIRECT**	**DIRECT**
ON-LINE	SELF-TUNING	Q-LEARNING
OFF-LINE	IDENTIFICATION + DESIGN	VRFT

«extreme points» of a range of intermediate arrangements

OFF-LINE VS ONLINE

OFF-LINE

C is tuned <u>once</u> (after observing S over a proper time window $t=N$)

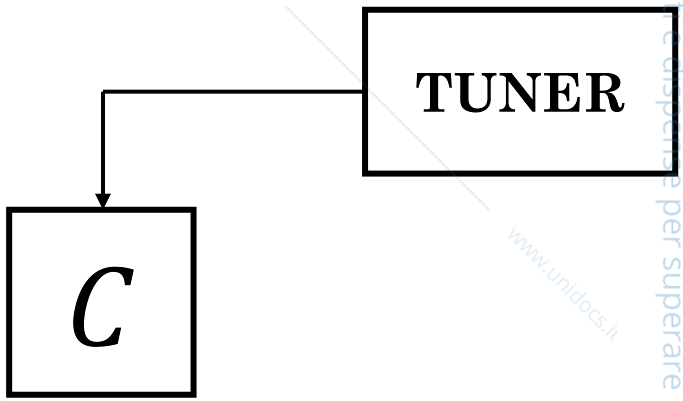


limitations:

- S must be t-invariant
- avoid poor performing controller up to $t=N$

ONLINE

C is tuned at every time (based on the available a-posteriori info)



- overcomes limitations of off schemes (but it is more complicated)
- can be used for both t-varyi and t-invariant S

<u>approach</u>: refer to the way the TUNER chooses C among $\boldsymbol{\mathcal{C}}$

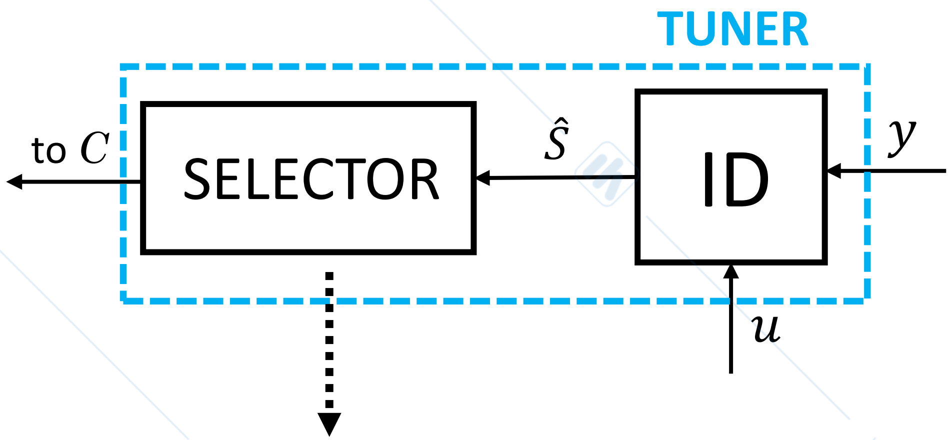
INDIRECT VS DIRECT

(1) <u>INDIRECT APPROACHES</u>
«curse of dimensionality»

$$\boxed{\textit{IDENTIFICATION ALGORITHM} + \textit{SELECTOR}}$$

TUNER first performs an identification of $S \Rightarrow$ estimate of $\hat{S}$ and then it decides C based on $\hat{S}$

<u>SELF-TUNING SCHEME</u>



choose C as if $\hat{S}$ was the true system we want to c

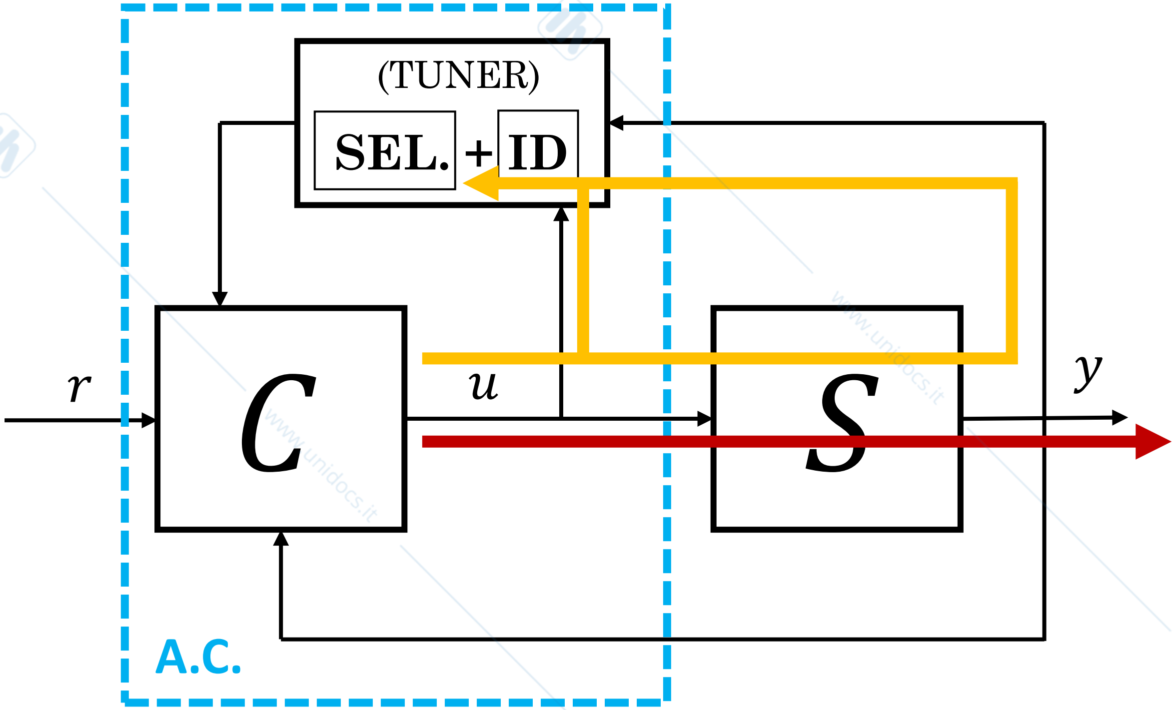
→ <u>CERTAINTY EQUIVALENCE PRINCIPLE</u> (CEP)

most used for its <u>simplicity</u> (very valuable): other more in principle high performing schemes are too complicate to any practical use

prescribes the «design» of C forgetting about the mismatch between the real and the identified mod (very large at the beginning)

$S \neq \hat{S}$ alwa

- **PRIMAL EFFECT**: the response by S as obtained given the input selected by the controller C (impact of the control action on the response of S)
- **DUAL EFFECT**: impact of control action u on the estimate $\hat{S}$ of $S \rightarrow$ indirectly related to the achievement of our control goals, because of the adaptation rul



EXPLOITATION VS EXPLORATION

u is used to make y as close as possible to the desired behavior (the more complex the task → the more the knowledge of S is required)

u is used to achieve the best possible identificat of S (u is used to excite and reduce the mismatc between S and $\hat{S}$)

CONFLICTUAL

control goal requires to keep S as steady-state as possible

we want to excite S as much as possible

In A.C. (S is uncertain) we have to acquire knowledge via exploration progressively, while enforcing the desired behavior via exploitation → optimal performance requires a proper trade-off (not achieved by self-tuning and CEP) between exploration and exploitation

focus more on

DUAL ADAPTIVE CONTROLLER: CEP is discarded and a better balance (the optimal balance) between exploration and exploitation is required (extremely complicated: very difficult to use in any practical problem)

ROBUST ADAPTIVE CONTROL: identification returns an estimate $\hat{S}$ of S along with an assessment of $S - \hat{S}$ (mismatch) that is used to select C so as to achieve a partial balance between exploration and exploitation

(2) DIRECT APPROACHES: *no identification* of S to decide C

CEP $\longrightarrow$ use $J(C,\hat{S})$ to determine a $\bar{C}$: $J(\bar{C},\hat{S}) \leq K$

real performance not satisfactory $\longrightarrow$ $S \neq \hat{S} \Rightarrow J(\bar{C},S) > K$ even if $J(\bar{C},\hat{S}) \leq K$

(perfomance index that counts)

for a good performance: $\underbrace{J(C,S) \approx J(C,\hat{S})}$

- unexplicit
- difficult to workout
- not pursued by existing

to achieve this we should be able to understand what are the features of S that are relevant for th determination of J and let Ŝ to estimate those feature only, otherwise we are identifying too much and the wealth of data is wasted in unessential parts to solve the control

---

use data to «identify» $J(C,S)$

$\downarrow$

$\hat{J}(C)$ performance index obtained from the data with the goal of making $\boxed{\hat{J}(C) \approx J(C,S)}$

Uncertain system $\boldsymbol{S}$

$\text{WN}(0, \lambda^2)$ additive disturbance

$$\boldsymbol{S}:\ y(t) = \underbrace{\varphi(t-1)^T}_{\text{regression vector}} \underbrace{\theta^0(t)}_{\text{parameter vector}} + e(t)$$

t-varying ARX system

operational representation (z^{-k} n-step delay operators)

delay between u and y

$$\boldsymbol{S}:\ A\left(z^{-1}, \theta^0(t)\right) y(t) = B\left(z^{-1}, \theta^0(t)\right) u(t-d) + e(t)$$

Hp:

- $d \geq 1$
- $b_0^0(t) \neq 0\ ,\ \forall t$
- $\theta^0(t)$: A and B are COPRIME $\forall t$ (no cancellations between them)
- observability and reachability (well-posed control problem)

- dependent variable → signals $\forall t > 0$

  ↓

  $y(t)$ (system output) → observable

- indipendent (tunable) variable : $u(t)$ control action

- indipendent uncertain quantities :
  - $y(-n+1)$ initialization
  - $e(t)$ additive disturbance
  - $a^0_{1-n}(t), b^0_{0-n}(t)$ t-varying parameters uncertain

  no need for adaptation (classical robust controllers

  motivates adaptation

Identification block objective of estimating the uncertain parameter on

↓

specific algorithm dictated by the a-priori info on $\boldsymbol{S}$

↓

map from $\underbrace{\text{data}}$ to an estimate of the parameter vector

$D^N = \{u(1), y(1), \dots, u(n), y(n)\}$ data set

Model Class (ARX)

$\mathrm{WN}(0, \sigma^2)$

estimated

$$\boldsymbol{M}(\theta):\ y(t) = \varphi(t-1)^T \underbrace{\theta(t)}_{\substack{\textbf{model}\\ \text{parameter}\\ \text{vector}}} + \xi(t)$$

stochastic model

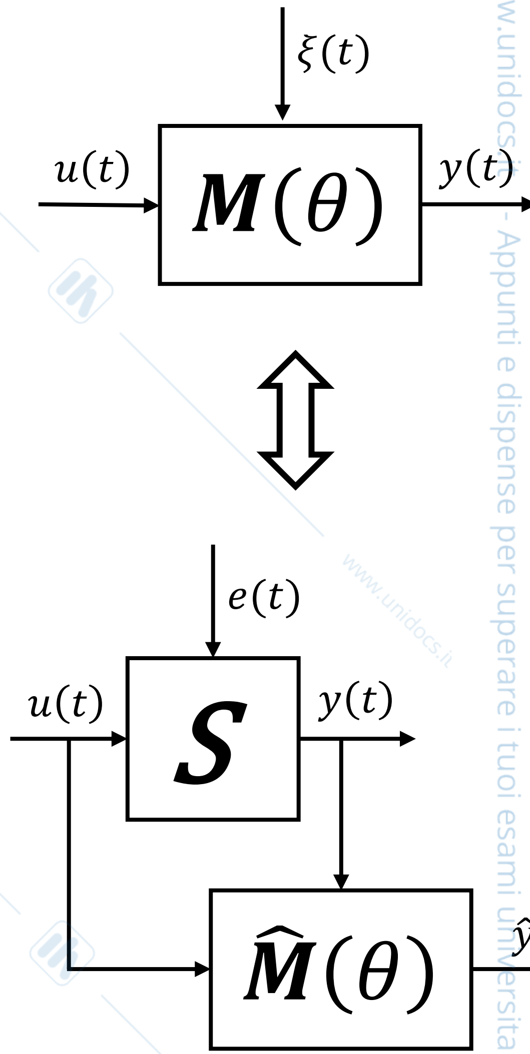
PEM identification (Prediction Error Minimization)

$$\boldsymbol{M}(\theta):\ y(t) = \underbrace{\varphi(t-1)^T\theta}_{\text{predictable}} + \underbrace{\cancel{\xi(t)}}_{\text{unpredictable}}$$

$\downarrow$ model in prediction form

$$\widehat{\boldsymbol{M}}(\theta): \hat{y}(t|t-1) = \varphi(t-1)^T\theta$$

$\hat{y}$ is a signal which can be compared with the measured y and the principle is to choose the model that has $\hat{y}$ as close as possible to y



<u>BEST</u> (unique) model parameter vector (Least Squares estimate)

$$\hat{\theta}_N = \underset{\theta}{\operatorname{argmin}}\, J_N(\theta)$$

$$= \underset{\theta}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \left(y(i) - \theta^T \varphi(i-1)\right)^2$$

quadratic positive cost function

↓

$$\frac{d}{d\theta} J_N(\theta) = 0$$ <u>condition</u> ← stationary points are <u>minimizers</u>

↓

$$\hat{\theta}_N = \underbrace{\left(\sum_{i=1}^{N} \varphi(i-1)\varphi(i-1)^T\right)^{-1}}_{\text{the inverse must be non-singular}} \sum_{i=1}^{N} \varphi(i-1)y(i$$

problem for A.C. → $\hat{\theta}_N$ need to be <u>well-defined</u> in all condition (A.C. has to be implemented as <u>standalone device</u> → must always return a well-defined value)

A.C. identification in close-loop $\longrightarrow$ data are collected when S operates with the controller → goal of the controller is to hide the dynamics o

A.C. singularity is often the case

↓

Modified LS

positive definite by assumption ⇒ non-singular

$$J_N(\theta) = \sum_{i=1}^{N} \underbrace{(y(i) - \varphi(i-1)^T\theta)^2}_{\text{prediction error (can be singular)}} + \underbrace{(\theta - \bar{\theta})^T S_0 (\theta - \bar{\theta})}_{\text{quadratic term (non-singular paraboloid)}}$$

↓

same calculations of standard L.S.

↓

$$\hat{\theta}_N = [S_0 + \textstyle\sum_{i=1}^{N} \varphi(i-1)\varphi(i-1)^T\,]^{-1}\,[S_0\bar{\theta} + \textstyle\sum_{i=1}^{N} \varphi(i-1)y(i)]$$

(parameter vector o

<u>Bayesian setting</u>: $(\theta - \bar{\theta})^T S_0 (\theta - \bar{\theta})$ can be interpreted as a-priori info on $\theta^0($

- $\bar{\theta}$ = <u>mean value</u> of $\theta^0(t_0)$
- S_0 = <u>dispersion</u> of $\theta^0(t_0)$ around $\bar{\theta}$

<u>a-priori distribution</u> characterizing the perception of uncertainty in S

- the <u>bigger</u> S_0 , the more importance is given to $\bar{\theta}$, the <u>slower</u> is the L.S. to mo away from $\bar{\theta}$ (in case $\bar{\theta}$ is wrong)
- the <u>smaller</u> S_0 , the closer to singularity (<u>numerical problems</u>)

Implementation

<u>self-tuning</u>: N, fixed number of data → **t**, estimate is continuously updated as time progresses (a new observation is selected at every time instant)

t is continuously increasing → data record is increasing with t → **<u>memory allocation problem</u>** (saturation of memory capacity)

↓

<u>Recursive LS</u> (RLS) <u>finite</u> memory allocation (no memory saturation)

$$\hat{\theta}_t = f(\hat{\theta}_{t-1}, y(t), \varphi(t-1))$$

identification algorithm is a <u>dynamical syst</u>
(is updated from the previously computati
→ recursion must <u>not</u> introduce any approximation)

$$\hat{\theta}_{t+1} = [\underbrace{S_0 + \sum_{i=1}^{t+1} \varphi(i-1)\varphi(i-1)^T}_{\text{information matrix } S(t+1)}]^{-1} [S_0\bar{\theta} + \sum_{i=1}^{t+1} \varphi(i-1)y(i)]$$

$$\hat{\theta}_{t+1} = \underbrace{\hat{\theta}_t}_{\text{previous estimate}} + \underbrace{S(t+1)^{-1}\varphi(t)}_{\text{weight}}\underbrace{[y(t+1) - \varphi(t)^T\hat{\theta}_t]}_{\text{innovation}}$$

recursive equation for LS estimate

innovation carried by $y(t+1)$ wrt the prediction we can make at time t

$$\varphi(t)^T\hat{\theta}_t = \hat{y}(t+1|t,\hat{\theta}_t)$$

RLS – I form

- finite memory allocation
- numerically instable ($S(t)$ may diverge)
- computationally demanding (inverse computation)

} not suitable for practical implementation

RLS – III form (computationally inexpensive and also numerically stable)

RLS – I : update $S(t)$, but then $S(t)^{-1}$ is only used to compute $\hat{\theta}_t$
→ update $S(t)^{-1}$ **directly**

Matrix Inversion Lemma

F, H, G, K matrices of proper dimension so that F, H, F + GHK are all invertible

$$(F + GHK)^{-1} = F^{-1} - F^{-1}G(H^{-1} + KFG)^{-1}KF^{-1}$$

update of $S(t)^{-1}$ via $\varphi(t)$ to obtain $S(t+1)^{-1}$
↓
computationally cheap (requires sum, products and inversion of a scalar)

$$S(t+1)^{-1} = S(t)^{-1} - \frac{S(t)^{-1}\varphi(t)\varphi(t)^T S(t)^{-1}}{1 + \varphi^T S(t)^{-1}\varphi(t)}$$

$$\varphi(t-1) = A\varphi(t-2) + B\begin{bmatrix} y(t-1) \\ u(t-d) \end{bmatrix}$$

<u>RLS is not able to track time-varying parameters</u>

asymptotically $\hat{\theta}_{t+1} = \hat{\theta}_t$ $\longrightarrow$ RLS is no more reactive to parameter changes as t increases

$$J_t(\theta) = \sum_{i=1}^{t} (y(i) - \varphi(i-1)^T\theta)^2 + (\theta - \bar{\theta})^T S_0 (\theta - \bar{\theta})$$

<u>all the information</u> (a-priori and a-posteriori for every time instant) are <u>weighted uniformly</u> even those data points that refers to a value of the S that is no more (<u>t-varying setup</u>)

$\downarrow$

<u>past data</u> should be <u>weighted less</u> than more recent data (<u>no more informative</u> about the current value of $\theta^0(t)$)

Least Squares with forgetting factor (μ-LS)

tuning parameter $\mu \in (0,1]$

$$J_t^{\mu}(\theta) = \sum_{i=1}^{t} \mu^{t-i}(y(i) - \varphi(i-1)^T\theta)^2 + \mu^t(\theta - \bar{\theta})^T S_0(\theta - \bar{\theta})$$

exponentially decaying weight

- $\mu \approx 1 \rightarrow$ more weight to past data $\longrightarrow$ the less erratic the estimate (the slower the convergence to $\theta^0(t)$)
- $\mu \approx 0 \rightarrow$ less weight to past data $\longrightarrow$ the more erratic the estimate (the faster the convergence)

disturbance case: the effect is not averaged and the a-posteriori info does not accumulate anymore

$\boldsymbol{\mu}$-LS estimate: $\hat{\theta}_t^{\mu} = argmin_{\theta}\, J_t^{\mu}(\theta)$

observations:

- $\mu = 1 \Rightarrow$ 1-LS ≡ LS
- $J_t^{\mu}(\theta)$ is still quadratic
- strictly needed for t-varying systems

$\boldsymbol{\mu}$ - RLS - I form

- $S(t)$ is as. stable ($\mu < 1$) (or constant)
- $S(t+1)^{-1} \neq 0 \Rightarrow \mu -$ RLS is always reactive to new info

Drawbacks of forgetting factor

as. stability cannot prevent $S(t)$ to get close to singularity

Blow-up phenomenon

$\hat{\theta}_t$ has a sudden variation (burst):

- instability (self-tuning) → $S(t)^{-1}$ is diverging along some elements
- numerical issues
- $\varphi(t)\varphi(t)^T$ singular (non informative identification experiments)

$\downarrow$ to avoid blow-up

t-varying μ

$$\mu(t) = \begin{cases} \mu \mid rcond\big(S(t)\big) \geq \alpha \\ 1 \mid rcond\big(S(t)\big) < \alpha \end{cases}$$

- $S(t)$ state of controller
- $rcond\big(S(t)\big)$ indicator of singularity of $S(t)$
- α least acceptable level of singularity

μ is tuned continuously based on $rcond\big(S(t)\big)$ → $S(t)$ is kept non-singular all the time → no blow-up

t-varying μ

- to avoid singularity conditions
- reduce the effect of wrong initialization in LS (prior info)

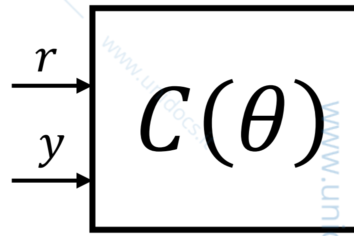
in $\mu - LS$ is vanishing fast ← $S(0) = S_0$ and $\hat{\theta}_0 = \bar{\theta}$ chosen at will by designer → may lead to slow converge (poor performance of self-

- erratic estimate $\hat{\theta}_t^{\mu}$ (noise effect is not canceled out)

$\mu = \mu(t) = \bar{\mu}$ for **small *t*** →
- reduce the effect of initialization when it is needed more
- few data points (A not able to counteract B)

$= 1$ for **big *t*** → many data capable to counteract B

«Basic» controller $C \in \boldsymbol{\mathcal{C}}$ in self-tuning



$C(\theta)$ → controller that has to be tuned depending on $\theta = \hat{\theta}_t$

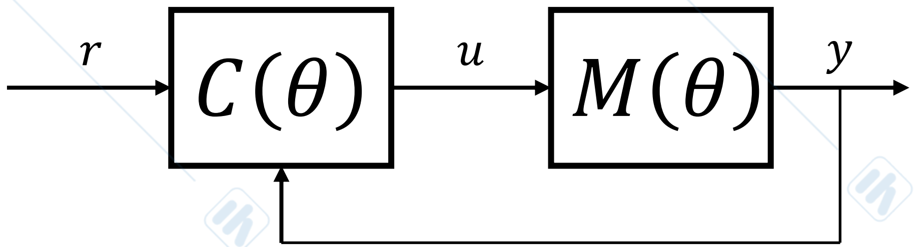
→ can be any linear proper controller whose parameters are functions of the model parameter vector θ

perform
achieved
nominal

$$\boldsymbol{C}(\theta): A_C(z,\theta)u(t) = B_C(z,\theta)y(t) + C_C(z,\theta)r(t)$$

«self-tuning is as good as the control scheme obtained by connecting $C(\theta)$ to M

$\sum(\theta,\theta)$ control scheme ($C(\theta)$ is feedback connected with $M(\theta)$)



we don't know which θ is suc
that self-tuning ≈ $\sum(\theta,\theta)$

we would like that $\overbrace{\sum(\theta,\theta)}^{\text{linear t-invariant}}$ is $\underbrace{\text{satisfactory}}$ for the largest set of θ values

it achieves the control goals (dictates the choice of $\boldsymbol{C}$) specified at the beginning

$$\Xi = \{\theta: \sum(\theta,\theta) \text{ is well-defined and satisfactory}\}$$

- in most control techniques Ξ is very large
- still it can't be $\Xi = \Theta$ (whole domain for θ)
  $\rightarrow$ intrinsic limitations (it is important to spot out what Ξ is)

Minimum Variance Control (MVC)

$$\boldsymbol{M}(\theta): A(z,\theta)y(t) = B(z,\theta)u(t-d) + \xi(t)$$

MV designs $u(t)$ so as to minimize the tracking error variance

$$\min_{u} t[y(t+d) - r(t)]^2 \longleftarrow$$

MV controller is obtained by solving this wrt to u

min is achieved when $\underbrace{\hat{y}(t+d|t)}_{\text{predictor of } y(t+d)} = r(t)$

$y(t+d) = r(t)$ is the best tracking possible (without being anticipative wi d-steps delay)

$\boldsymbol{C}(\theta)$: $B(z,\theta)E(z,\theta)u(t) = r(t) - F(z,\theta)y(t)$ <u>equation for MV co</u>

$\Xi = \{\theta: \sum(\theta,\theta)$ is <u>satisfatcory</u>$\}$

<u>Hp</u>: as.stability is always with the goals of control design

$\Xi^* = \{\theta: \sum(\theta,\theta)$ is <u>as.stable</u>$\}$

$\Rightarrow \Xi \subseteq \Xi^*$

$M(\theta)$ must be minimum phase in order to $\sum(\theta,\theta)$ be as.stable

$\Xi \subseteq \{\theta: \sum(\theta,\theta)$ is <u>minimum phase</u>$\}$

main limitation of MVC

if the system and its inverse are <u>causal</u> and <u>stable</u>

Variants: **Generalized MV Control** (GMVC) → does not require the minimum phase condition

intrinsic limitations:
- observability
- reachability

→ otherwise we are missing the degrees of freedom to operate with $M(\theta)$ → the control problem is ill-posed

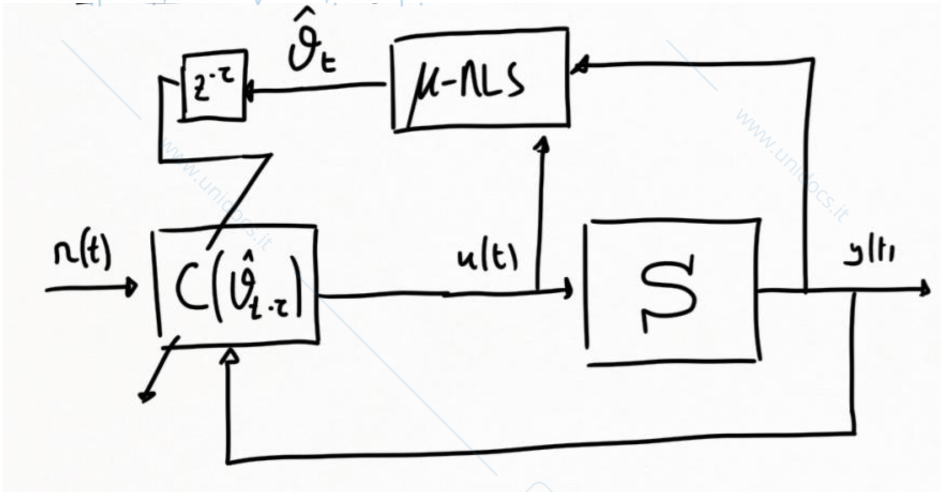
↓

in general $\Xi \subseteq \{\theta : \sum(\theta, \theta) \text{ is reachable and observable}\}$ largest possible
→ pole placement is a technique that make $\Xi = \{\theta : \sum(\theta, \theta)$ is reachable and ob

↑

by choosing K, L, N properly, based on $F(\theta), G(\theta), H(\theta)$ for all θ, we can set eigenvalues (desired poles in closed-loop) and static gain of the control scheme at will such that $M(\theta)$ is observable and reachable

Self-tuning



$$\Sigma(\theta_t^0, \hat{\theta}_{t-\tau})$$

$\boldsymbol{S} = \boldsymbol{M}(\theta_t^0)$ $\quad$ $\boldsymbol{C}(\hat{\theta}_{t-\tau})$ $\quad$ with $\hat{\theta}_{t-\tau}$ LS estimate constructed based on data collected fro

$\theta_t^0 = \theta^0 \ (\mu = 1) \leftrightarrow \boldsymbol{S}$ t-invariant — complete theoretical results showing the robustness of $\Sigma(\theta_t^0, \hat{\theta}_{t-\tau})$

$\theta_t^0 \ (\mu < 1) \leftrightarrow \boldsymbol{S}$ t-varying — results are only partial and research is quite open but we have empirical evidence

Convergence of RLS

- $\hat{\theta}_t$ always converge (asymptotically) to a finite value, which will be denoted by
- $\hat{\theta}_t$ cannot diverge or oscillate on a limit cycle
- no assumption on the operation of $\boldsymbol{S}$ is made (convergence arises even when are not informative)

estimation error $\tilde{\theta}_t = \hat{\theta}_t - \theta^0$

convergence $\tilde{\theta}_t \leftrightarrow$ convergence $\hat{\theta}_t \leftrightarrow$ convergence $S(t)^{-1}$

$S(t+1) \geq S(t) \geq S_0 > 0 \; \forall t$ S_0 is positive definite (to avoid singulari

Linear Algebra th: $A \geq B > 0 \leftrightarrow B^{-1} \geq A^{-1} > 0$ $S(t)^{-1}$ is "bounded" abo

$S(t)^{-1} \rightarrow_{t\rightarrow\infty} \rightarrow \bar{S}^{-1}$ fixed matrix

$= 0 \rightarrow S$ is perfectly identified asymptotically ($S(t)$ diverges along all possib

in general $S \neq 0$

Self-optimality

«output» of $\sum(\theta^0, \hat{\theta}_t)$: $z(t) = \binom{y(t)}{u(t)}$ «output» of $\sum(\hat{\theta}_\infty, \hat{\theta}_\infty)$: $z(t) = \binom{y_{i,\infty}(t)}{u_{i,\infty}(t)}$

A self-tuning $\sum(\theta^0, \hat{\theta}_t)$ is said to be self-optimal if $\lim_{t\to\infty} z(t) - z_{i,\infty}(t) = 0$

$\downarrow$

self-tuning $\sum(\theta^0, \hat{\theta}_t)$ is asymptotically behaving as as.imaginary scheme $\sum(\hat{\theta}_\infty,$

$\swarrow$

behavior of $\sum(\hat{\theta}_\infty, \hat{\theta}_\infty)$ depends on Ξ (depends on $\boldsymbol{\mathcal{C}}$)

satisfactory if and only if $\hat{\theta}_\infty \in \Xi$

\+

self-optimality of $\sum(\theta^0, \hat{\theta}_t)$

$\sum(\theta^0, \hat{\theta}_t)$ is satisfactory (even though θ^0 is unknown) [1]

Th: fundamental in self-tuning

if $\hat{\theta}_\infty \in \Xi$, then $\sum\left(\theta^0, \hat{\theta}_t\right)$ is self-optimal [2]

[1] + [2]

$\hat{\theta}_\infty \in \Xi \rightarrow \sum\left(\theta^0, \hat{\theta}_t\right)$ (self-tuning) behaves satisfactorily as for the given control specification
(even though θ^0 is not known, Ξ is typically a "large" set and $\hat{\theta}_\infty \in \Xi$ happens rob

$\sum(\theta, \theta)$ is a standard linear control scheme, where controller is connected to model based on which controller is designed

$\theta \notin \Xi$ are exceptions due to instrinsic limitations

Literature of self-tuning → characterization of the robustness of the condition $\hat{\theta}_\infty$

[1]. $\hat{\theta}_\infty$ asymptotic LS estimate , it depends on θ^0 , experiment $(r(t), e(t))$ and initialization $\bar{\theta}, S_0$

result: irrespective of θ^0 and experiment , if $\bar{\theta}$ is chosen at random, then $\hat{\theta}_\infty \in \Xi$ with **probability 1**

the set is large
chance of missi
asymptotically
perturbation is
null)

we can inspect wheter $\hat{\theta}_t$ is approaching Ξ^C (complement)

[2]. Ξ is known to user (it is a byproduct of our choice $\boldsymbol{C}$)

LS can be modified (while maintaing all the features of LS) so that when $\hat{\theta}_t$ is too c to Ξ^C , a perturbation is introduced so as to eventually guarantee $\hat{\theta}_\infty \in \Xi$ always

$\sum(\theta^0, \hat{\theta}_t)$ is obtained as a <u>feedback perturbation</u> $p(t)$ (non-linear) over $\sum(\hat{\theta}_t,$ (imaginary system)

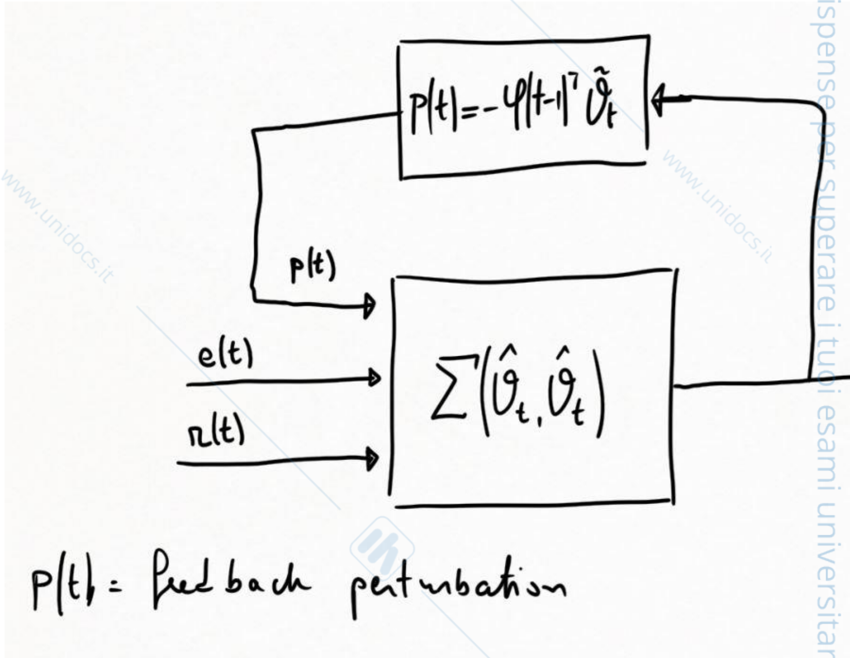
output of $\sum(\hat{\theta}_t, \hat{\theta}_t)$ when $p(t) \xrightarrow[t\to\infty]{} 0$ (asymptotically vanishing)

=

output of $\sum(\theta^0, \hat{\theta}_t) \to z(t) \xrightarrow[t\to\infty]{} z_{i,\infty}(t)$ output of $\sum(\hat{\theta}_\infty, \hat{\theta}_\infty)$

$$p(t) = -\varphi(t-1)^T \tilde{\theta}_t \longrightarrow$$ LS estimation error

regression vector



for components of $\tilde{\theta}_t$ such that:

- $\tilde{\theta}_\infty \neq 0$ (there isn't info to identify θ^0), then $\varphi(t-1) \to 0$ along the same components (info is vanishing)
- $\tilde{\theta}_\infty = 0$, then $\varphi(t-1)$ is bounded (feedback)

In order to $\sum\left(\theta^0, \hat{\theta}_t\right)$ be satisfactory, it is not required that $\hat{\theta}_\infty = \theta^0$

$\hat{\theta}_\infty \neq \theta^0$ means that data are poorly informative and $p(t) \xrightarrow[t\to\infty]{} 0$ means that knowledge of θ^0 is not needed to accomplish the control task

Whenever the control task requires to know θ^0 , then the operation of S becomes informative and LS must be consistent $\hat{\theta}_\infty = \theta^0$ ($\tilde{\theta}_\infty = 0$)

Non-linear systems

$$\mathrm{WN}(\bar{e},\lambda^2)$$

known to the user

$$y(t) = \boldsymbol{f}(y(t-1),\dots,y(t-n),u(t-d),\dots,u(t-d-m),e(t))$$

non linear

pb: tracking of the set-point $r(t)$

$$e(t) = \underbrace{\bar{e}}_{\text{nominal value}} + \underbrace{\delta e(t)}_{\text{deviation}}$$

Hp: $r(t) = \bar{r} + \delta r(t)$

the controller can be designed by means of linearization in a neighborhood of an operating condition (deviations must be small enough)

controller → control of the nominal value

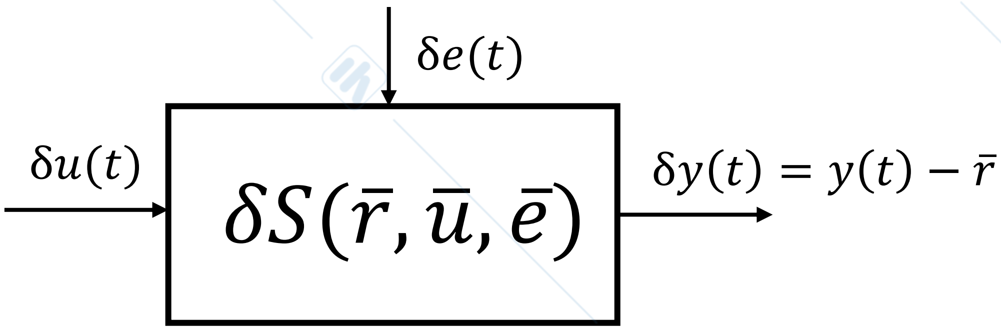
controller → control of deviation

$$u(t) = \bar{u} + \delta u(t)$$

let $\bar{u}$ such that in nominal conditions (disturbance is set to $\bar{e}$ only)
$u(t) = \bar{u}\ \ \forall t$ leads to $y(t) = \bar{r}\ \ \forall t$

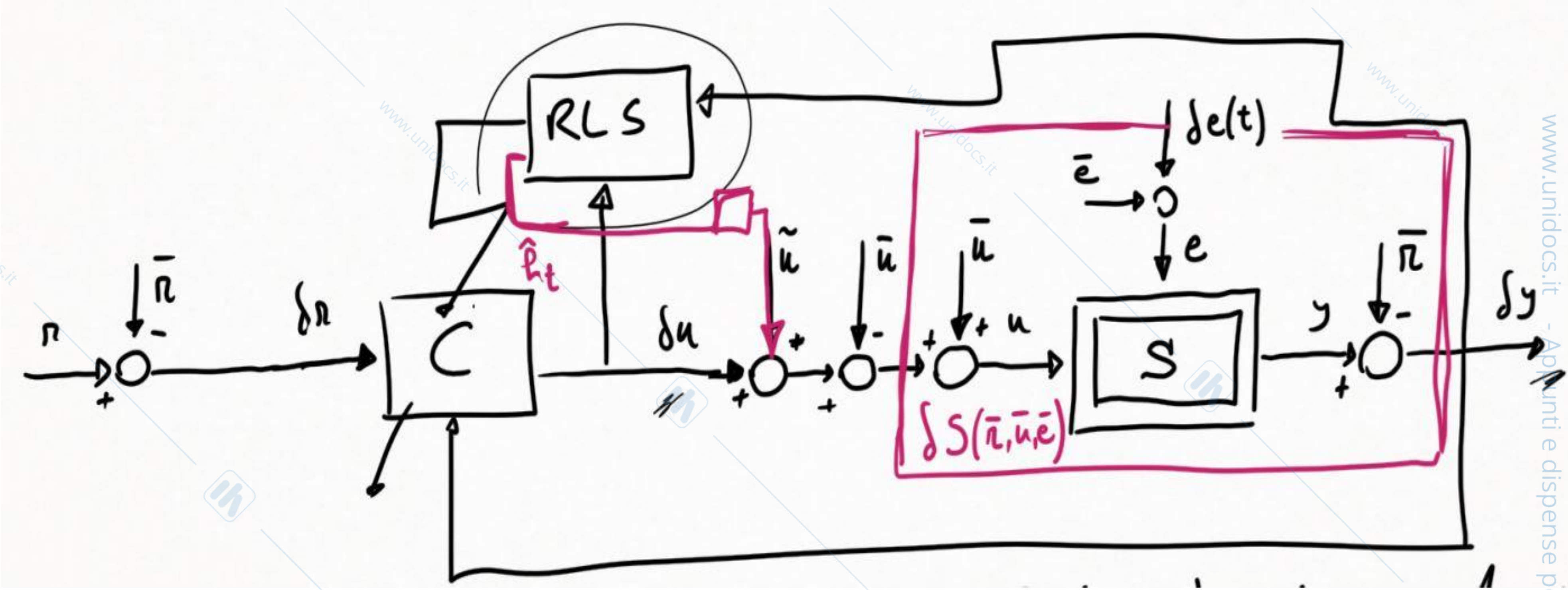
$\bar{r} = \boldsymbol{f}(\bar{r}, \bar{u}, \bar{e}) \longleftarrow$ $\bar{u}$ = solution to this equation

Linearized uncertain ARX system (parametric uncertainty)



$\bar{e}$ is unknown → $\bar{u}$ is unknown
actual nominal control action
is $\tilde{u} \neq \bar{u}$

self-tuning can be used to design the controller for $\delta u(t)$



<u>constant distburance on input</u> h^0 → unknown parameter that need to be accoun by RLS to ensure <u>consistency</u> $\hat{h}_t$

the effect of $\tilde{u} - \bar{u}$ can be tamed by the basic controller but can also use $\hat{h}_t$ to re the error $\tilde{u} - \bar{u}$ (t-varying μ-RLS)

Virtual Reference Feedback Tuning (VRFT)

- off-line
- direct (essential feature)

setup: uncertain linear system (discrete t.f. form) we want to control by means of linear system

$$y(t) = P(z)u(t)$$

$P(z)$ is unknown

the controller must be selected among a parametric class of linear controllers $\boldsymbol{C} = \{C(z,\theta), \theta \in \Theta\}$

θ = controller parameter

control specifications: **Model Reference** → $M(z)$ = a desired transfer function for the behavior of th closed-loop control scheme

poles, zeros, gain of the desired closed-loop

r-to-y t.f. = $\dfrac{P(z)C(z,\theta)}{1+P(z)C(z,\theta)}$

$\boxed{H_2 \text{ norm}}$

$\uparrow$

control goal: to make the t.f. «as close as possible» to $M(z)$

Model Reference performance index

$$J_{MR}(\theta) = \left\| \left(\frac{P(z)C(z,\theta)}{1+P(z)C(z,\theta)} - M(z) \right) W(z) \right\|_2^2 =$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{P(e^{j\omega})C(e^{j\omega},\theta)}{1+P(e^{j\omega})C(e^{j\omega},\theta)} - M(e^{j\omega}) \right|^2 \underbrace{\left|W(e^{j\omega})\right|^2}_{\text{frequency weight}} d\omega$$

integ
freq.
smal
frequ

frequency response: $z = e^{j\omega}\ , \omega = [-\pi, \pi]$

$P(z)$ is unknown $\rightarrow$ missing the knowledge to compute $J_{MR}(\theta)$

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

calls for <u>adaptation</u> $\rightarrow C(z, \theta)$ design based on data

<u>experiment on the plant</u>: $\bar{u}, \bar{y}$ signals measured in the experiment
$\rightarrow$ data set D^N

<u>Virtual Reference</u>: $\bar{\boldsymbol{r}}_{\boldsymbol{v}}(\boldsymbol{i}) \quad i = 1, \ldots, N$ signal constructed from D^N

<u>off-line computed</u> $\leftarrow \bar{r}_v(i) = M(z)^{-1}\bar{y}(i) \quad i = 1, \ldots, N$

$\bar{r}_v(i)$ is a <u>particular set-point</u> (reference) such that the measured output $\bar{y}(i) \;\; i = 1, \ldots, N$ corresponds to the desired close-loop response according to the model reference $M(z)$

$M(z)r(t)$ desired response for the close-loop for $r(t)$

desired response for $\bar{r}(t) \rightarrow M(z)\bar{r}_v(t) = \bar{y}(i)$

<u>virtual error</u>: $\bar{e}_v(i) = \bar{r}_v(i) - \bar{y}(i) \quad i = 1, \ldots, N$

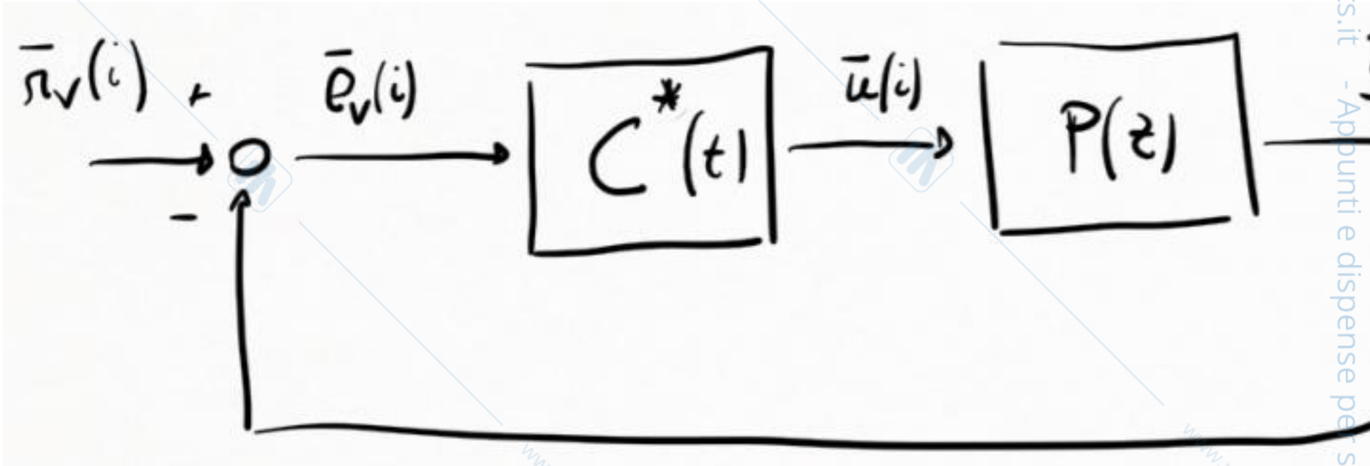
a good Model Reference controller is such that when fed by $\bar{e}_v(i)$ (virtual error) it returns $\bar{u}(i)$ (input in the experiment) as output

good controller:

$$\frac{PC^*}{1+PC^*}\bar{r}_v(i) = \bar{y}(i)$$

$\updownarrow$

M



$$\boxed{C^*\bar{e}_v(i) = \bar{u}(i)} \quad [*]$$

$\downarrow$

<u>necessary</u> condition for a controller to be good (sufficient if $\bar{e}_v(i)$ is variegated – experiment exciting)

VRFT control design

we impose $[*]$ to design our controller within $\boldsymbol{\mathcal{C}} = \{C(z,\theta): \theta \in \Theta\}$
$\rightarrow$ response of $C(z,\theta)$ to $\bar{e}_v(i)$ is as close as possible to $\bar{u}(i) \quad i = 1, \dots, N$

$$J_{VR}(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left(\bar{u}(i) - C(z,\theta)\bar{e}_v(i)\right)^2$$

$$\hat{\theta}_{VR} = \underset{\theta}{\operatorname{argmin}}\, J_{VR}(\theta)$$

[1] **direct approach**: no attempt to reconstruct $P(t)$, data are only used to enforce the controller a property that is required to achieve control goals

[2] $J_{VR}(\theta)$ has the same structure of cost in PEM identification (VRFT ≈ "identification of controller")

[3] generalizations:

generic metric

- $J_{VR}(\theta) = \frac{1}{N}\sum_{i=1}^{N} d(\bar{u}_i, C(z,\theta)\bar{e}_v(i))$
- prefiltering $L(z) \rightarrow \overline{(u,e,y)}_L = L(z)\overline{(u,e_v,y)}$ $\quad L(z) = 1$ original setup

$$J_{VR}(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left(\bar{u}_L(i) - C(z,\theta)\bar{e}_L(i)\right)^2$$

$L(z)$ degree of freedom to tune VRFT method so as to attain MR control desig as much as possible

remark: with linearly parametrized controllers, we have the same structure of LS identification → the minimizer (VRFT controller) can be computed explicitel and with extremely low computational effort

Frequency interpretation of PEM identification

(frequency characterization of PEM identification cost function)

$$\boldsymbol{S}:\ y(t) = G^0(z)u(t) + H^0(z)e(t) \sim \boxed{\mathrm{WN}(0,\lambda^2)}$$

linear stochastic system

rational t.f. (unknown)

$u(t) \perp e(\tau)\ \forall t,\tau$
stochastic
uncorrelated
(experiment)

Hp:

- G^0 and H^0 are as.stable
- $u(t)$ stationary stochastic process → $y(t)$ s.s.p.
- $H^0(z)$ → canonical ARMA (not restrictive)

Model class: $\mathcal{M} = \{M(\theta), \theta \in \Theta\}$ (linear stochastic system)

$$\boldsymbol{M}(\theta):\ y(t) = G(z,\theta)u(t) + H(z,\theta)\xi(t) \sim \boxed{\mathrm{WN}(0,\sigma^2)}$$

linear t.f. whose parameters are parametrized in θ
(black-box: θ = coeffs of NUM and DEN of G and H)

in general $G(z,\theta$
and $\mathrm{H}(z,\theta) \neq H$

Hp: $M(\theta)$ is such that $H(z,\theta)\xi(t)$ is canonical with $|zeros| < 1$
(condition under which prediction of SSP can be developed)

PEM: a good model of $\boldsymbol{S}$ is a model capable of predicting well future
$M(\theta) \rightarrow \widehat{M}(\theta)$ (from stochastic model to 1-step predictor models)

$$\hat{J}_N(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left(y(i) - \hat{y}(i|i-1,\theta)\right)^2$$

index of model $M(\theta)$ prediction capabilities
(depends on the experiment)

$$\hat{\theta}_N = \underset{\theta}{argmin}\, \hat{J}_N($$

PEM identification best mo

extremely difficult to characterize (finite sample identification results)

$\hat{J}_N(\theta, D_N)$ D_N: family of functions , one for each possible realization of the data set (experiments)

asymptotic theory of PEM identification: asymptotically as N increases, the family of functions shrink onto a unique curve

$$\hat{J}_N(\theta) \xrightarrow[N \to \infty]{} \bar{J}(\theta)$$

$$\hat{\theta}_N \xrightarrow[N \to \infty]{} \theta^*$$

almost surely (for all possible data sets , all users are asymptotically experiencing the same results – irrespective of the experiment)

we assume N large enough so that differences because of the specific experiments become negligible ($\hat{J}_N(\theta) \approx \bar{J}(\theta)$ and $\hat{\theta}_N \approx \theta^*$)

frequency interpretation is a characterization for $\bar{J}(\theta)$ ($\approx \hat{J}_N(\theta)$ for large N)

$$\bar{J}(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\left|G^0(e^{j\omega}) - G(e^{j\omega},\theta)\right|^2}{|H(e^{j\omega},\theta)|^2} \cdot \underbrace{\Phi_u(\omega)}_{\text{spectrum of } u(t)} d\omega + \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\left|H^0(e^{j\omega})\right|^2}{|H(e^{j\omega},\theta)|^2} \cdot \underbrace{\Phi_e(}_{=\lambda^2 \text{ spect}}$$

fundamental result: that reveals the contribution of the mismatch $G^0(z) - G(z,\theta$
and between H^0 and $H(0)$
(understand how the u-to-y model error is dislocated frequency by frequency)

prefiltering:

$$\bar{J}(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|G^0 - G(\theta)|^2}{|H(\theta)|^2} \cdot |L|^2 \Phi_u d\omega + \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|H^0|^2}{|H(\theta)|^2} \cdot |L|^2 \Phi_e(\omega) d\omega$$

L = degree of freedom to tune PEM identification

[1] Full parametrization (never happens, but error can be considered negligible in many app

$$\exists \theta^0 \in \Theta : G(\theta^0) = G^0 \text{ and } H(\theta^0) = H^0$$

$$\bar{J}(\theta) = \underset{(\geq 0\ \forall\theta)}{A(\theta)} + \underset{(\geq \lambda^2\ \forall\theta)}{B(\theta)} \quad \rightarrow \bar{J}(\theta^0 = \theta^*) = 0 + \lambda^2$$

θ^0 simultaneously achieve the minimum of both $A(\theta)$ and $B(\theta)$

$$M(\hat{\theta}_N) \xrightarrow[N\to\infty]{} M(\theta^*) = M(\theta^0) = S \qquad \boxed{\textbf{CONSISTENCY}}$$

condition for informativeness of data points $\Phi_u \neq 0 \rightarrow$ minimizer is unique

[2] Partial full-parametrization

$$\exists \theta^0 \in \Theta : G(\theta^0) = G^0 \text{ but } \underset{\text{wrong noise model}}{H(\theta) \neq H^0} \;\; \forall\theta$$

[2.1] <u>**Indipendent parametrization**</u> (distinct parameters)

$$\theta = \begin{bmatrix} \theta_G \\ \theta_H \end{bmatrix} \qquad G(\theta) = G(\theta_G) \qquad H(\theta) = H(\theta_H)$$

$\exists \theta_G^0 : G\left(\theta_G^0\right) = G^0$ but $H(\theta_H) \neq H^0$

$$\theta^* = \begin{bmatrix} \theta_G^0 \\ \theta_H^* \end{bmatrix} \qquad H(\theta_H^*) \neq H^0$$

PARTIAL CONSISTENCY

[2.2] <u>**Indipendent parametrization**</u> (common parameters)

θ^0 minimizer of $A \neq \hat{\theta}$ minimizer of B

$$\bar{J}(\theta^*) \leq \bar{J}(\theta) \ \ \forall\theta \qquad \theta^* \neq \theta^0 \text{ and } \theta^* \neq \hat{\theta}$$

minimum achieved as a <u>compromise</u> of the two

$G\left(\hat{\theta}_N\right) \xrightarrow[N\to\infty]{} G(\theta^*) \neq G\left(\theta_G^0\right) = G^0$ and $H(\theta^*) \neq H^0$

NO PARTIAL CONSISTE

[3] Under parametrization

$\forall\theta : G(\theta) \neq G^0$ and $H(\theta) \neq H^0$

NO CONSISTENCY

whatever θ^* is $G(\theta^*) \neq G^0$ and $H(\theta) \neq H^0$

Still in case [2.2] and [3]

mismatch between the system and model u-to-y t.f.
$|G^0 - G(\theta)|^2$ concerns in the construction of $\bar{J}(\theta)$
with weight $\frac{\Phi_u}{|H(\theta)|^2}$ (small → big mismatch)

Hp: high enough signal-t ratio (neglect B term)

weight at the optimum → $\frac{\Phi_u}{|H(\theta^*)|^2} \approx \frac{\Phi_u}{\left|H(\hat{\theta}_N)\right|^2}$

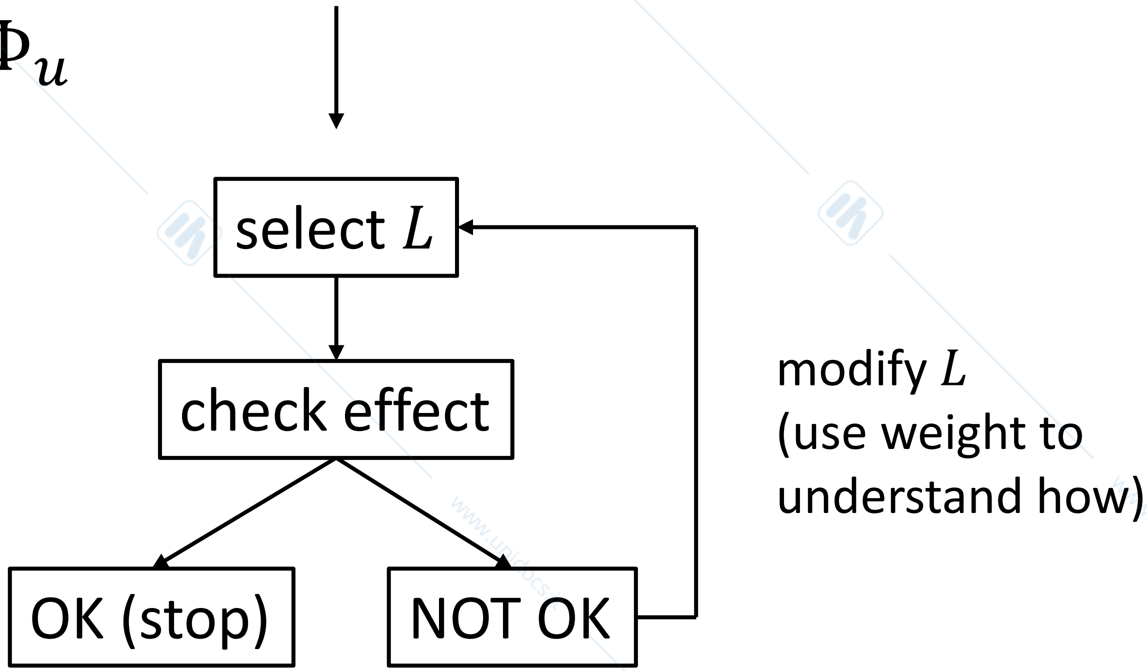
(evolution of mismatch f by f)

with <u>prefiltering</u>:

freq. weight $\dfrac{|L|^2}{|H(\theta^*)|^2}\Phi_u$

degree of freedom to tune (<u>iteratively</u>) the mismatch

<u>a-posteriori revealed</u> (inspect weight to understand mismatch)



<u>identification for control</u>: a tool to drive the information in the data to the identification of part of $\boldsymbol{S}$ that is of interest (low-high freq.)

PEM

$\hat{\theta}_N^{OE}$ $\quad \left|G^0(z) - G\left(z, \hat{\theta}^{OE}\right)\right|^2_{z=e^{j\omega}}$

freq. weight = 1

uniformly distributed all over the freq. range

$\hat{\theta}_N^{ARX}$ $\quad \left|G^0(z) - G\left(z, \hat{\theta}^{ARX}\right)\right|^2_{z=e^{j\omega}}$ weighted by $\left|A\left(e^{j\omega}, \hat{\theta}^{A}\right.\right.$

prefiltering

$$\left|G^0(z) - G\left(z, \hat{\theta}_L{}^{ARX}\right)\right|^2_{z=e^{j\omega}}$$

weighted by

$$|L|^2 \left|A\left(\hat{\theta}_L{}^{ARX}\right)\right|^2$$

close-loop identification

full-parametrization → consistency
(in all other cases we have unconsistency – no partial consistency)

mismatch $|G^0 - G(\theta^*)|^2$

weighted by

$$\frac{|C|^2}{|1 + CG^0|^2 |H(\theta^*)|^2}$$

unknown (even a-posteriori)

$G(\theta^*)$

approximation (not the best)

IDENT 1: $y(t) = G(\theta^*)u(t) + H(\theta^*$

IDENT 2: $u(t) = W(\beta^*)r(t) + Z(\beta^*$

$$W(\beta^*) \approx \frac{C}{1 + CG^0}$$

$$\frac{|C|^2}{|1 + CG^0|^2 |H(\theta^*)|^2} \approx |W(\beta^*)|^2 \frac{1}{|H(\theta^*)|^2}$$

→ approximation for the freq
for the mismatch in IDENT

VRFT performance (in comparison with MR)

ideal controller: $C^0(z)$ digital system (t.f.)

$$\frac{P(z)C^0(z)}{1+P(z)C^0(z)} = M(z) \leftrightarrow C^0 = \frac{M(z)}{P(z)(1-M(z))}$$

C^0 is always well-
as long as $M(z)$ ≠
(mild requiremen

$C^0(z)$ is introduced for theoretical purposes only:

- $C^0(z)$ cannot be computed in practice (missing the knowledge of $P(z)$)
- $C^0(z)$ does not belong to $\boldsymbol{\mathcal{C}}$ in general ($C^0(z)$ may even be anticipative → not implementable)

$C^0(z)$ is used to rewrite J_{MR} and J_{VR} so as to enhance comparison

<u>Hp</u> (mild → $\bar{u}$ is user chosen): $\bar{u}(i)$ is a ssp realization

<u>as. theory of PEM identification</u>

$J_{VR}(\theta) \xrightarrow[N\to\infty]{} \bar{J}_{VR}(\theta)$ uniformly in θ (irrespective of the realization of $\bar{u}(i)$)

<u>Hp</u>: N is large enough → <u>mismatch is negligible</u> $J_{VR}(\theta) \approx \bar{J}_{VR}(\theta)$

$$\bar{J}_{VR}(\theta) = E\left[\left(L(z)\bar{u}(t) - C(z,\theta)L(z)\bar{e}_v(t)\right)^2\right] =$$

$$= \mathrm{E}\left[\left(\frac{L(z)}{M(z)} \cdot \frac{P(z)(C^0(z) - C(z,\theta))}{1 + P(z)C^0(z)} \cdot \bar{u}(t)\right)^2\right] = E[\bar{x}(t)^2]$$

$$\bar{J}_{VR}(\theta) = E[\bar{x}(t)^2] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|P|^2|C^0 - C(\theta)|^2}{|1 + PC^0|^2} \cdot \boxed{\frac{|L|^2}{|M|^2}} \Phi_u d\omega$$

$$J_{MR}(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|P|^2|C^0 - C(\theta)|^2}{|1 + PC^0|^2} \cdot \boxed{\frac{|W|^2}{|1 + PC(\theta)|^2}} d\omega$$

$\exists \theta^0 : C(z, \theta^0) = C^0(z) \longrightarrow$ $\mathcal{C}$ large enough to contain the ideal controller (never in practice)

$J_{MR}(\theta^0) = 0 = \bar{J}_{VR}(\theta^0) \longrightarrow$ MR and VRFT have the <u>same minimizer</u> θ^0

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari
www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

$\forall\theta : C(z,\theta) \neq C^0(z)$ approximation is <u>not negligible</u>

$$\Rightarrow J_{MR}(\theta) \neq \bar{J}_{VR}(\theta) \Rightarrow \underset{\theta}{\operatorname{argmin}} J_{MR}(\theta) \neq \underset{\theta}{\operatorname{argmin}} \bar{J}_{VR}(\theta)$$

<u>idea</u>: use prefilter $L(z)$ to make $\bar{J}_{VR}(\theta)$ as close as possible to $J_{MR}(\theta)$

<u>naive approach</u>

$$\frac{|L|^2}{|M|^2}\Phi_u = \frac{|W|^2}{|1+PC(\theta)|^2} \Rightarrow J_{MR}(\theta) = \bar{J}_{VR}(\theta)$$

$$|L|^2 = \frac{|W|^2|M|^2}{|1+PC(\theta)|^2\Phi_u} \Rightarrow L(z) = \frac{W(z)M(z)}{\big(1+P(z)C(z,\theta)\big)G(z)}$$

1. $P(z)$ is unknown
2. $C(z,\theta)$ depends on θ

→ <u>prefiltering cannot be parametric</u> in θ

spectral factori

$G(z) : |G|^2 =$

[2] use C^0 (unknown) in place of $C(\theta)$

$\downarrow$

however approximation is good at least in a neighborhood of θ^{MR} (optimal controller for MR)

$$\theta^{MR} = \underset{\theta}{\operatorname{argmin}} J_{MR}(\theta)$$

θ^{MR} is the point where $\frac{1}{1+PC^0}$ and $\frac{1}{1+PC(\theta)}$ are as close as possible

$\rightarrow \theta^{MR}$ is the point where the difference between $J_{MR}(\theta)$ and $\bar{J}_{VR}(\theta)$ is minimal

$\theta^*_{VR} \approx \theta^{MR} \rightarrow \theta^*_{VR}$ (VRFT controller) has MR performance close to θ^{MR} (MR contr

[1] $L(z) = \frac{W(z)M(z)}{(1+P(z)C^0(z))G(z)} = \frac{W(z)M(z)(1-M(z))}{G(z)}$ everything known and implementable

«optimal» prefilter given the predent state knowledge

Output is affected by an additive disturbance $(\bar{u}(t) \perp d(t))$

$$\bar{J}_{VR}(\theta) = \underset{\geq 0}{A(\theta)} + \underset{\geq 0}{B(\theta)} \rightarrow$$ bias because of disturbance

even if $\boldsymbol{\mathcal{C}}$ is large enough so that $\exists\theta^0 : C(\theta^0) = C^0$
$\rightarrow A(\theta^0) = 0\,, B(\theta^0) \neq 0$

θ⁰ not minimizer for
(min achieved for $C($

$\bar{J}_{VR}(\theta)$ compromises between A and B
$\rightarrow \underset{\theta}{\operatorname{argmin}}\, \bar{J}_{VR}(\theta) = \theta^* \neq \theta^0$

$C(\theta^*) \neq C^0 \rightarrow$ even if $C^0 \in \boldsymbol{\mathcal{C}}$
(C^0 and θ^0 makes $J_{MR}(\theta) = 0$) $\longrightarrow$ VRFT <u>not consistent</u> with MR

<u>VRFT</u> scheme as proposed cannot get rid of <u>bias</u>:
- $L(z)$ can be used to make $A(\theta) \approx J_{MR}(\theta)$
- $L(z)$ is not able to make $A(\theta) + B(\theta) \approx J_{MR}(\theta)$

change VRFT scheme

PEM identification is not enough to make VRFT working with disturbance

→ use other identification of controller approaches (IV identification)

Instrumental variable

1. $C^0(z) \in \boldsymbol{\mathcal{C}} \quad (\exists \theta^0 : C(z,\theta^0) = C^0(z))$

2. $C(z,\theta) = \theta^T \beta(z) \;\; \theta = [\theta_1 \dots \theta_n]^T \;\; \beta(z) = [\beta_1(z) \dots \beta_n(z)]$

PEM VRFT: $J_{VR}(\theta) = \frac{1}{N}\sum_{i=1}^{N}(\bar{u}_L(i) - \theta^T \varphi_L(i))^2$

1.+2. $\exists \theta^0 : C^0(z) = {\theta^0}^T \beta(z)$

$\boxed{d = 0}$ $\quad \bar{u}_L(i) = C^0(z)\bar{e}_L = \varphi_L(i)^T \theta^0$

no disturbance

$$\hat{\theta}_{VR} = \underset{\theta}{\operatorname{argmin}}\, J_{VR}(\theta) = \theta^0 \longrightarrow \boxed{\textbf{CONSISTENCY OF VRFT}}$$

$\boxed{d \neq 0}$ $$\bar{u}_L(i) = \varphi_L(i)^T \theta^0 + z(t)$$ $z(t)$ colored noise (even when $z(t)$ is w

$$\hat{\theta}_{VR} \xrightarrow[N \to \infty]{} \theta^0 + \text{bias} \longrightarrow$$ VRFT not consistent with MR

$\downarrow$

φ_L and z are correlated
(both depends on d at
past time indexes)

IV VRFT (new definition of $\hat{\theta}_{VR}$)

$$\hat{\theta}_{VR}^{IV} = \left[\sum_{1}^{N} \zeta(i)\varphi_L(i)^T\right]^{-1} \times \sum_{1}^{N} \zeta(i)\,\bar{u}_L(i)$$

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari
www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

$\zeta(t)$ is the instrumental variable and must be a vectorial signal such that:

- $E[\zeta(t)\varphi_L(t)^T] \neq 0$ non-singular with prob. 1 ($\hat{\theta}_{VR}^{IV}$ is well-posed)
- $E[\zeta(t)d(\tau)] = 0 \quad \forall t,\tau \quad \zeta \perp d$ (IV is uncorrelated with disturbance)

$\hat{\theta}_{VR} \xrightarrow[N\to\infty]{} \theta^0 \longrightarrow$ **CONSISTENCY WITH MR** (asymptotically)

to obtain $\zeta(t)$ → **REPEATED EXPERIMENT**

correlation in disturbance at differen
time instants tend to vanish with lag

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari
www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

Reinforcement Learning (RL)

Markov Chain Decision Process (MDP)

$$x_{t+1} = f(x_t, u_t, w_t)$$

state-space discrete-time stochastic system

w_t: exogenous disturbance stochastic indipendent

$$\forall t \begin{cases} x_t \in S_x = \{x^1, \dots, x^{n_x}\} \\ u_t \in S_u = \{u^1, \dots, u^{n_u}\} \end{cases}$$

superscript=index of possible values

finite number of possible states and control actions

stochastic system + quantized state and control input → transition probabilities

$$p\left(x^k \middle| x^i, u^j\right) = prob\left\{x_{t+1} = x^k \middle| x_t = x^i \wedge u_t = u^j\right\}$$

$p(x^k | x^i, u^j)$ $\forall x^k, x^i, u^j$ completely specifies the evolution of a MCDP because
$x_{t+1} = f(x_t, u_t, w_t)$ has the so called <u>**MARKOV PROPERTY**</u>:

$$prob\{x_{t+1} = x^k | x_t = x^i, x_{t-1} = x^r, x_{t-2} = x^e, \ldots \wedge u_t = u^j\}$$

$$= prob\{x_{t+1} = x^k | x_t = x^i \wedge u_t = u^j\}$$

state in the past is <u>irrilevant</u> for the evolution of x_{t+1} once x_t in the present is kn
(intrinsic because the notion of state)

$$\sum_{k=1}^{n_k} p(x^k | x^i, u^j) = 1 \qquad\qquad p(x^k | x^i, u^j) \neq 0 \text{ (in general)}$$

$p(* \mid *, u^j)$ one matrix $\forall u^j \in S_u$

- diagrams = chains
- nodes = state values
- edges = transition probs

x, u are continuous variables → <u>quantization</u> generates (approximately) a MCDP

DECISION PROBLEM

$$\min_{\{u_t\}} J(x, \{u_t\})$$

sequence of u_t

x_t (evolution of state of MCDP) is stochastic

→ sequence of input is deterministic (poor performing)

→ feedback is needed (input adapted or some information on the state)

Hp: state x_t is observable and available at time t

$$u_t = \pi_t(x_t)$$

in general t-dependent

$x_t = x^i, u_t = u^i \rightarrow x_{t+1} = x^k$ generates:

- loss $L(x^i, u^j)$ (control)
- reward $R(x^i, u^j)$ (machine learning)

→ random variable ($L(x^i, u^j)$ is defi $f(x_{t+1})$ which is random)

MCDP is starting from $x_0 = x \in S_x$ (initial state) → stochastic evolution of MCD

sequence of inputs → stochastic sequence of losses

performance is associated with average of losses: $E[L(x_0, u_0)] \dots E[(L(x_t, u_t)$

Discounted average loss

$$J(x, \{u_t\}) = \sum_{t=0}^{+\infty} \gamma^t E[L(x_t, u_t)]$$

$\gamma \in (0,1)$ discount factor (exponenti decaying importance to loss as t incr

Richard Bellman

value function (optimal discounted total average loss)

$$V^*(x) = \min_{u_0, \{u_t\}_{t=1}^{+\infty}} J(x, \{u_t\}) =$$

$$= \min_{\{u_t\}} \left[E[L(x, u_0)] + \gamma \sum_{i=1}^{n_x} p\left(x^i \middle| x, u_0\right) \cdot E\left[\sum_{t=1}^{+\infty} \gamma^{t-1} L\left(x_t^i, u_t\right) \right] \right]$$

the role of $x_0 = x$ and u_0 in the construction of $V^*(x)$ is now explicitly indicate

$x_t^i, t \geq 1$ = evolution of the MCDP when $x_1 = x^i$ $i = 1, \dots, n_x$

Bellman optimality principle (dynamic programming)

determine first the optimal strategy from $t = 1$ forward
(for all possible values taken by x_1 according to what is seen: u_t adapted to the s
$\{u_t\}_{t=1}^{+\infty}$ is chosen so as to behave optimally for each value of x
use u_0 to optimize the first step and the subsequent evolution, taking into accou
transition probabilities and the fact that from $t = 1$ forward we will be optimal

$$V^*(x) = \min_{u_0} \left[E[L(x, u_0)] + \gamma \sum_{i=1}^{n_x} p\left(x^i \middle| x, u_0 \right) \cdot \min_{\{u_t\}_{t=1}^{+\infty}} \sum_{t=1}^{+\infty} \gamma^{t-1} E\left[L\left(x_t^i, u_t\right)\right] \right]$$

decide u_0 so as to trade-off between expected loss for $t = 0$ and the optimal ave
behavior from $t = 1$ formward

MCDP is t-invariant

optimal choice of u_t is given a t-invariant feedback (unique policy applied at eve

$$\boxed{u_t = \pi(x_t)}$$

to achieve optimality it is enough to solve the problem of the selection o

$$V^*(x) = \min_{u_0} \left[E[L(x, u_0)] + \gamma \sum_{i=1}^{n_x} p(x^i | x, u_0) \cdot V^*(x^i) \right]$$

Bellman functional equation (always admits a unique solution)

$$\begin{cases} V^*(x^1) \\ V^*(x^2) \\ \cdot \\ \cdot \\ \cdot \\ V^*(x^{n_x}) \end{cases}$$

system of n_x equations in n_x unknowns → non-linear becaus of $\min_{u_0}$ (piecewise linear)

one can solve for all possible combinations of u_0 for each equation and then a-posteriori check which one satisfies

$$u_0^* = \pi^*(x) = \underset{u_0}{\operatorname{argmin}} \left[E[L(x, u_0)] + \gamma \sum_{i=1}^{n_x} p(x^i | x, u_0) \cdot V^*(x^i) \right]$$

<u>optimal selection</u> of u_0 as a function of x

to compute π^* we need to know:

1. $E[L(x, u_0)] \rightarrow$ mechanism through which losses are generated in the system under study
2. $p(x^i | x, u_0) \rightarrow$ mechanism through which the system under study evolves

<u>uncertainty</u>
mechanisms
are not know
to the user

<u>**Q-learning**</u> (on-line and direct)

observations of MCDP are used to retrieve from data the optimal control policy <u>reconstructing</u> $p(x^i | x, u_0)$ and $E[L(x, u_0)]$ (leveraging the Bellman optimal pri

Q^* function $\quad Q^*: S_x \times S_u \to \mathbb{R}$

<u>unknown</u> ($Q^*(x,u)$ not directly avalable)

$$Q^*(x,u) = E[L(x,u_0)] + \gamma \sum_{i=1}^{n_x} p(x^i | x, u_0) \cdot V^*(x^i)$$

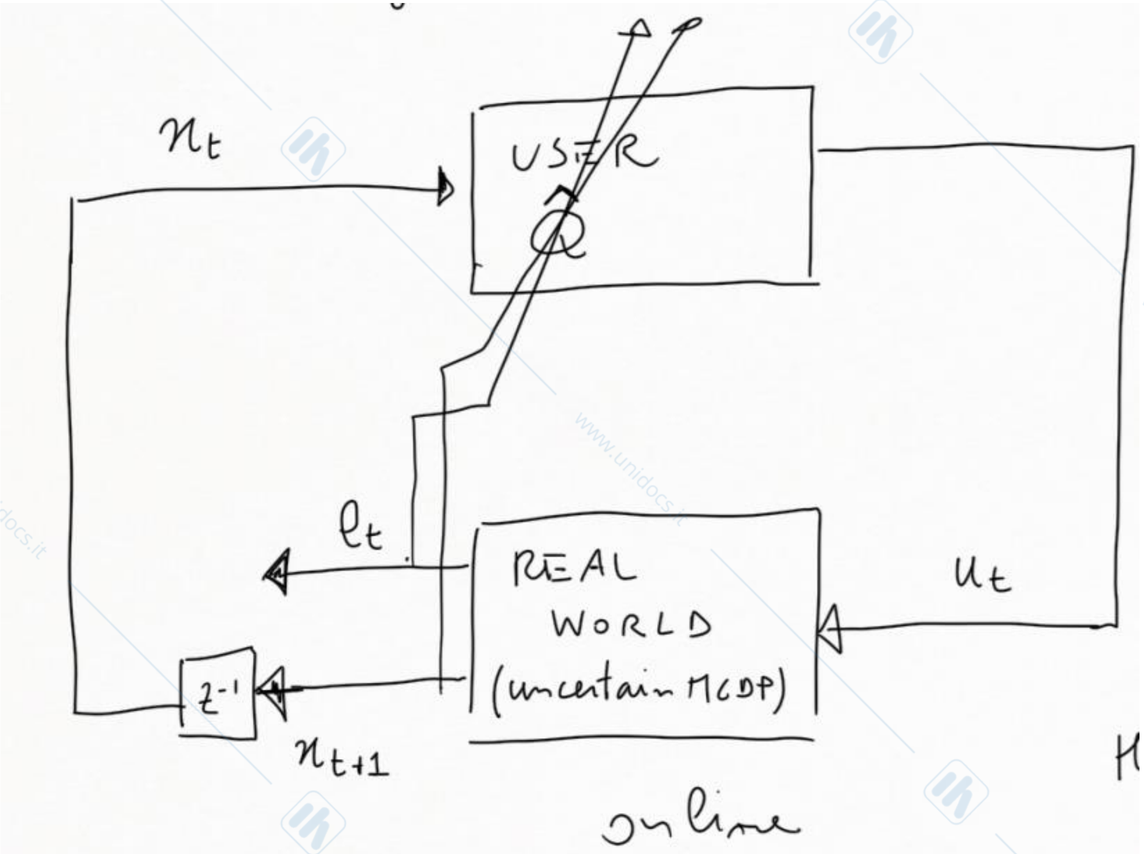
$Q^*(x,u)$ is the total discounted average loss when $x_0 = x$, u_0 is taken as u (without optimizing) and then $\{u_t\}_{t=1}^{+\infty}$ are optimally chosen ($u_t = \pi^*(x_t)$ for t

$$V^*(x) = \min_u Q^*(x,u) \to \pi^*(x) = \underset{u_0}{\operatorname{argmin}}\, Q^*(x,u)$$

optimizing u in Q^* function gives the <u>optimal solution</u>

Q-learning → learn Q^* from data to find the optimal policy:

- user at time t is given x_t and based on the current knowledge decides u_t
- u_t is given as input to the system we are dealing with which generates x_{t+1} (evolution of state) and an <u>instance</u> of loss l_t ($L(x_t, u_t)$)
- x_{t+1} and l_t are used to update the user's current knowledge ($\hat{Q}$)



realizations of state and loss are <u>observable</u>

Q-learning algorithm

$$Q_t(x,u) : S_x \times S_u \to \mathbb{R}$$

approximation of Q^* at time t (representable through a matrix)

Q_t is updated from Q_{t-1} based on the fact that the state was x_t and input was selected as u_t and also that x_{t+1} and l_t were observed

$$Q_t(x,u) = \begin{cases} (1-\alpha_t)Q_{t-1}(x,u) + \alpha_t\big(l_t + \gamma V_{t-1}(x_{t+1})\big) & x = x_t\,, u = u_t \\ Q_{t+1}(x,u) & x \neq x_t\,, u \neq u_t \end{cases}$$

α_t t-varying parameter
→ <u>MIXING COEFFICIENT</u>
(consistent cases only)

$$V_{t-1}(x) = \min_u Q_{t-1}(x,u)$$

approximated value function based on Q_{t-1}

Th: $T(x, u) = \{t : x_t = x , u_t = u\}$ set of t indexes

If α_t : $\sum_{t \in T(x,u)} \alpha_t^2 < +\infty$, $\sum_{t \in T(x,u)} \alpha_t = +\infty$ $\forall x \in S_x , \forall u \in S_u$

then with probability 1 $Q_t(x, u) \xrightarrow[t \to \infty]{} Q^*(x, u)$ $\forall x, u$

$\downarrow$

characterizes the condition (informativeness of the experiment) so as to asymptotically retrieve the optimal solution to the decision problem from experiments (requires enough exploration)

Th. holds irrespective of $Q_0(x, u)$