

Dipartimento di Ingegneria Gestionale, dell'Informazione e della Produzione



ADAPTIVE LEARNING, ESTIMATION AND SUPERVISION OF DYNAMICAL SYSTEMS (ALES)

Case study 1: Virtual Reference Feedback Tuning (VRFT)

Master Degree in COMPUTER ENGINEERING

Data Science and Data Engineering Curriculum

SPEAKER

Prof. Mirko Mazzoleni

PLACE

University of Bergamo

Syllabus

1. Recursive and adaptive identification

- 1.1 Recursive ARX estimation (RLS)
- 1.2 Least Mean Squares (LMS)
- 1.3 Instrumental Variables (IV)

2. Closed-loop identification

3. Subspace and MIMO identification

- 3.1 Singular Value Decomposition
- 3.2 Impulse data: Ho-Kalman, Kung algorithms
- 3.3 Generic I/O data: the MOESP algorithm

4. Supervision of dynamical systems

- 4.1 Introduction to fault diagnosis
- 4.2 Model-based fault diagnosis
- 4.3 Parity space approaches
- 4.4 Observer-based approaches
- 4.5 Signal-based fault diagnosis
- 4.6 Knowledge-based fault diagnosis

CASE STUDIES

- Virtual Reference Feedback Tuning
- Nuclear particles classification
- Leak detection in an industrial valve
- Bearing fault identification

Outline

1. Application study: Virtual Reference Feedback Tuning (VRFT)

2. Defining the control problem

3. Basic idea

4. VRFT problem solution

5. Example

Outline

1. Application study: Virtual Reference Feedback Tuning (VRFT)

2. Defining the control problem

3. Basic idea

4. VRFT problem solution

5. Example

Virtual Reference Feedback Tuning (VRFT)

The VRFT is a **direct method** for designing controllers. It operates using a **«batch»** of input-output (open loop) data collected from the system

Indirect control design methods

- 1. Perform experiments on the open-loop system
- 2. Identify a model of the system
- 3. Design the controller based on identified model

Traditional ****model-based control design*** paradigm

Direct control design methods

- 1. Perform experiments on the open-loop system
- 2. Design directly the controller («identify the controller»)

Output Output Output Data-drivencontrol design paradigm



Ingredients of the VRFT method

1. Unknown SISO linear system y(t) = G(z)u(t)

G(z) is **not known** (and we do not want to identify it)

2. Family of linear parametric 1-DOF controllers

$$R(z; \boldsymbol{\theta}) = \boldsymbol{\beta}^{\mathsf{T}}(z)\boldsymbol{\theta} \qquad \boldsymbol{\beta}(z) = [\beta_1(z) \ \beta_2(z) \dots \beta_d(z)]^{\mathsf{T}} \qquad \boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \dots \theta_d]^{\mathsf{T}}$$

$$\boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \dots \theta_d]^{\mathsf{T}}$$

Examples:

$$\boldsymbol{\beta}(z) = \begin{bmatrix} 1 & z^{-1} & \dots & z^{-n} \end{bmatrix}^{\mathsf{T}}$$

$$\boldsymbol{\beta}(z) = \begin{bmatrix} 1 & z^{-1} & \dots & z^{-n} \end{bmatrix}^{\mathsf{T}} \qquad \qquad \Box \qquad \qquad R(z;\boldsymbol{\theta}) = \theta_1 + \theta_2 z^{-1} + \dots + \theta_n z^{-n}$$

$$\beta(z) = \begin{bmatrix} 1 & \frac{1}{1-z^{-1}} & 1-z^{-1} \end{bmatrix}^{\mathsf{T}} \qquad \Box > \qquad R(z; \theta) = \theta_1 + \theta_2 \frac{1}{1-z^{-1}} + \theta_3 (1-z^{-1})$$

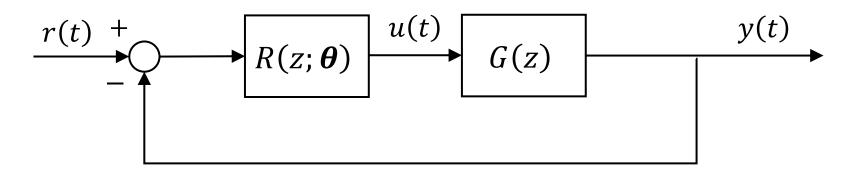
PID controller

Ingredients of the VRFT method

3. Model-reference control specification

Let M(z) a reference model for the desired closed-loop behaviour. The aim is to design the controller $R(z; \theta)$ so that

$$M(z) \approx \frac{G(z)R(z; \boldsymbol{\theta})}{1 + G(z)R(z; \boldsymbol{\theta})}$$



Outline

1. Application study: Virtual Reference Feedback Tuning (VRFT)

2. Defining the control problem

3. Basic idea

4. VRFT problem solution

5. Example

Defining the control problem

What we would like to minimize is the (possibly weighted) closed-loop mismatch between the reference model and the attained behaviour given the controller $R(z; \theta)$

$$J_{\text{MR}}(\boldsymbol{\theta}) = \left\| \left(\frac{G(z)R(z;\boldsymbol{\theta})}{1 + G(z)R(z;\boldsymbol{\theta})} - M(z) \right) \cdot W(z) \right\|_{2}^{2}$$

Model-reference cost

This cost **cannot be computed** since G(z) is not known!

We have to find another cost function (different but similar) that can be minimized, using a set of open-loop system measurements

Outline

1. Application study: Virtual Reference Feedback Tuning (VRFT)

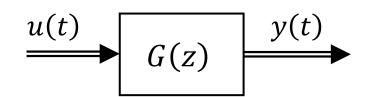
2. Defining the control problem

3. Basic idea

4. VRFT problem solution

5. Example

1. Perform an **open-loop experiment** on the system, and collect the measurements $\mathcal{D} = \{u(t), y(t)\}_{t=1}^{N}$



2. IF I was in closed-loop, and IF the closed loop worked perfectly, then the output y(t) would have been such that $y(t) = M(z)\bar{r}(t)$

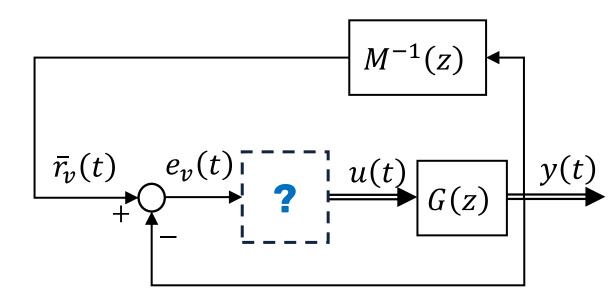
where $\bar{r}(t)$ is the **reference signal** that generated the measurement y(t). The signal $\bar{r}(t)$ is **«virtual»** since it does not exist, but it can be computed as

$$\bar{r}(t) = M^{-1}(z)y(t)$$

3. It is then possible to define the **virtual tracking error** $e_v(t)$ of this closed-loop (which does not exist)

$$e_v(t) = \bar{r}(t) - y(t)$$

4. The controller that grants this closed-loop to exist must be such that, when fed with the tracking error $e_v(t)$, would provide the input u(t) that generated y(t)



Thus, the problem is to identify the controller $R(z; \theta)$ from $\mathcal{D} = \{u(t), y(t)\}_{t=1}^N$ by minimizing

$$J_{\text{VR}}^{N}(\boldsymbol{\theta}) = \frac{1}{N} \sum_{t=1}^{N} \left(u(t) - R(z; \boldsymbol{\theta}) e_{v}(t) \right)^{2}$$

Virtual-reference cost

Remark 1:

M(z) is typically a strictly proper transfer function, so that $M^{-1}(z)$ will result into a **non-causal** transfer function, i.e. computing $\bar{r}(t)$ requires y(t+1), y(t+2), ...

This is not a problem, since we are working **off-line**, in a **batch** manner. So, future samples of $y(\cdot)$ are already available

The reference model should be $M(z) \neq 1$, otherwise

$$e_v(t) = \bar{r}(t) - y(t) = (M^{-1}(z) - 1)y(t) = 0$$

Remark 2:

Using a controller which **is linear in the parameters**, the solution can be computed using least squares. In fact, the linear in the parameter controller is described by

$$R(z; \boldsymbol{\theta}) = \boldsymbol{\beta}^{\mathsf{T}}(z)\boldsymbol{\theta}$$

Thus

$$R(z; \boldsymbol{\theta}) e_v(t) = \boldsymbol{\beta}^{\mathsf{T}}(z) \boldsymbol{\theta} e_v(t) = \boldsymbol{\beta}^{\mathsf{T}}(z) e_v(t) \boldsymbol{\theta} = \boldsymbol{\varphi}^{\mathsf{T}}(t) \boldsymbol{\theta}$$

where

$$\boldsymbol{\varphi}(t) = [\beta_1(z)e_v(t) \quad \beta_2(z)e_v(t) \quad \dots \quad \beta_d(z)e_v(t)]^{\mathsf{T}}$$

Then, the virtual reference cost function $J_{\rm VR}(\theta)$ can be rewritten as

$$J_{\text{VR}}^{N}(\boldsymbol{\theta}) = \frac{1}{N} \sum_{t=1}^{N} \left(u(t) - R(z; \boldsymbol{\theta}) e_{v}(t) \right)^{2} = \frac{1}{N} \sum_{t=1}^{N} (u(t) - \boldsymbol{\varphi}^{\mathsf{T}}(t) \boldsymbol{\theta})^{2}$$

So that the solution is

$$\widehat{\boldsymbol{\theta}}_{\text{LS}} = \left(\sum_{t=1}^{N} \boldsymbol{\varphi}(t) \boldsymbol{\varphi}^{\mathsf{T}}(t)\right)^{-1} \cdot \sum_{t=1}^{N} \boldsymbol{\varphi}(t) u(t)$$

Remark 3:

We modified the original problem of minimizing $J_{\rm MR}(\theta)$ into the minimization of $J_{\rm VR}^N(\theta)$. The two cost functions may not have the same minimum!

In which measure the **two minima differ?** To answer this question, the $J_{VR}^{N}(\theta)$ function is slightly modified by considering a pre-filtered version of input and virtual error signals

$$e_L(t) = L(z)e_v(t)$$

$$u_L(t) = L(z)u(t)$$

$$u_L(t) = L(z)u(t)$$



$$J_{\text{VR}}^{N}(\boldsymbol{\theta}) = \frac{1}{N} \sum_{t=1}^{N} \left(u_{L}(t) - R(z; \boldsymbol{\theta}) e_{L}(t) \right)^{2} = \frac{1}{N} \sum_{t=1}^{N} (u_{L}(t) - \boldsymbol{\varphi}_{L}^{\mathsf{T}}(t) \boldsymbol{\theta})^{2}$$

$$\boldsymbol{\varphi}_L^{\mathsf{T}}(t) = \boldsymbol{\beta}^{\mathsf{T}}(z)e_L(t)$$

Outline

1. Application study: Virtual Reference Feedback Tuning (VRFT)

2. Defining the control problem

3. Basic idea

4. VRFT problem solution

5. Example

In general, $J_{MR}(\theta)$ and $J_{VR}^N(\theta)$ are different and they have different minima. However, it is possible to **choose the pre-filter** L(z) so that the **two minima coincide**

Consider the model-reference cost $J_{MR}(\theta)$: the (squared) \mathcal{H}_2 norm can be written as

$$J_{\mathrm{MR}}(\boldsymbol{\theta}) = \left\| \left(\frac{G(z)R(z;\boldsymbol{\theta})}{1 + G(z)R(z;\boldsymbol{\theta})} - M(z) \right) \cdot W(z) \right\|_{2}^{2}$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{G(e^{j\omega})R(e^{j\omega};\boldsymbol{\theta})}{1 + G(e^{j\omega})R(e^{j\omega},\boldsymbol{\theta})} - M(e^{j\omega}) \right|^{2} \cdot \left| W(e^{j\omega}) \right|^{2} d\omega$$

The reference model M(z) can be expressed as

$$M(z) = \frac{G(z)R_0(z)}{1 + G(z)R_0(z)}$$

where $R_0(z)$ is the controller that **exactly solves** the model-matching problem. Notice that $R_0(z; \theta^0)$ could not belong to the model set $R(z; \theta)$

Substituting this definition of M(z) in $J_{MR}(\theta)$, we get (dropping $e^{j\omega}$ for simplicity)

$$J_{\text{MR}}(\boldsymbol{\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|G|^2 |W|^2}{|1 + GR(\boldsymbol{\theta})|^2} \cdot \frac{|R(\boldsymbol{\theta}) - R_0|^2}{|1 + GR_0|^2} d\omega$$

Consider now the **virtual-reference cost** $J_{VR}^{N}(\theta)$

$$J_{\text{VR}}^{N}(\boldsymbol{\theta}) = \frac{1}{N} \sum_{t=1}^{N} (u_{L}(t) - R(z; \boldsymbol{\theta}) e_{L}(t))^{2}$$

If u(t) and y(t) are realizations of **stationary** and **ergodic processes**, also $\bar{r}(t)$, $e_v(t)$, $e_L(t)$, $u_L(t)$ are stationary, since computed by linear filtering. Then

$$J_{\text{VR}}^{N}(\boldsymbol{\theta}) \rightarrow J_{\text{VR}}(\boldsymbol{\theta}) = \mathbb{E}\left[\left(u_{L}(t) - R(z; \boldsymbol{\theta})e_{L}(t)\right)^{2}\right]$$

so that $J_{VR}(\theta)$ is the **variance** of the stationary process $u_L(t) - R(z; \theta)e_L(t)$

We can also write

$$u_L(t) - R(z; \boldsymbol{\theta}) e_L(t) = L(z) u(t) - L(z) R(z; \boldsymbol{\theta}) e_v(t) = L(z) u(t) - L(z) R(z; \boldsymbol{\theta}) [M^{-1}(z) - 1] y(t)$$

$$= L(z)u(t) - L(z)R(z; \theta)[M^{-1}(z) - 1]G(z)u(t) = L(z)\left[1 - R(z; \theta)G(z)\left(\frac{1}{M(z)} - 1\right)\right]u(t)$$

$$= \frac{L(z)}{M(z)} \left[M(z) - R(z; \boldsymbol{\theta}) G(z) \left(1 - M(z) \right) G(z) \right] u(t)$$

$$= \frac{L(z)}{M(z)} \left[\frac{G(z)R_0(z)}{1 + G(z)R_0(z)} - R(z; \boldsymbol{\theta})G(z) \left(1 - \frac{G(z)R_0(z)}{1 + G(z)R_0(z)} \right) G(z) \right] u(t)$$

$$=\frac{L(z)}{M(z)}\frac{G(z)\big(R_0(z)-R(z;\boldsymbol{\theta})\big)}{1+G(z)R_0(z)}u(t)$$

The asymptotic virtual-reference cost $J_{VR}(\theta)$ can thus be written as:

$$J_{VR}(\boldsymbol{\theta}) = \mathbb{E}\left[\left(u_L(t) - R(z; \boldsymbol{\theta})e_L(t)\right)^2\right] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|L|^2}{|M|^2} \frac{|G|^2}{|1 + GR_0|^2} |R_0 - R(\boldsymbol{\theta})|^2 \cdot \Phi_{uu} d\omega$$

where $\Phi_{uu}(z)$ is the **power spectral density** of the input u(t)

$$J_{\text{MR}}(\boldsymbol{\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|G|^2 |W|^2}{|1 + GR(\boldsymbol{\theta})|^2} \cdot \frac{|R(\boldsymbol{\theta}) - R_0|^2}{|1 + GR_0|^2} d\omega$$

Model-reference cost

$$J_{VR}(\boldsymbol{\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|L|^2}{|M|^2} \frac{|G|^2}{|1 + GR_0|^2} |R_0 - R(\boldsymbol{\theta})|^2 \cdot \Phi_{uu} d\omega$$

Virtual-reference cost

Result 1: $R_0(z) \in R(z; \boldsymbol{\theta})$

If $R_0(z) \in R(z; \theta)$, then it exists θ^0 s.t. $J_{MR}(\theta^0) = J_{VR}(\theta^0) = 0$, so that the VRFT approach is able to attain the **optimal controller** $R_0(z)$ that achieves M(z)

Result 2: $R_0(z) \notin R(z; \boldsymbol{\theta})$

If $R_0(z) \notin R(z; \theta)$, it is possible to choose L(z) so that $J_{MR}(\theta^0)$ and $J_{VR}(\theta^0)$ coincide. Recall

$$J_{\text{MR}}(\boldsymbol{\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|G|^2 |W|^2}{|1 + GR(\boldsymbol{\theta})|^2} \cdot \frac{|R(\boldsymbol{\theta}) - R_0|^2}{|1 + GR_0|^2} d\omega \qquad J_{\text{VR}}(\boldsymbol{\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|G|^2 |L|^2}{|M|^2} \cdot \frac{|R_0 - R(\boldsymbol{\theta})|^2}{|1 + GR_0|^2} \cdot \Phi_{uu} d\omega$$

$$L(z) = \frac{M(z)W(z)}{1 + G(z)R(z; \boldsymbol{\theta})} \cdot \frac{1}{U(z)}$$

By choosing
$$L(z)$$
 so that $L(z) = \frac{M(z)W(z)}{1 + G(z)R(z;\theta)} \cdot \frac{1}{U(z)}$ $|L|^2 = \frac{|M|^2|W|^2}{|1 + GR(\theta)|^2} \cdot \frac{1}{\Phi_{uu}}$

it results that $J_{MR}(\theta) = J_{VR}(\theta)$, where U(z) is a spectral canonical factor of $\Phi_{uu}(\omega)$, s.t. $u(t) = U(z)\xi(t), \ \xi(t) \sim WN(0,1),$

Observation:

The **optimal filter**

$$L(z) = \frac{M(z)W(z)}{1 + G(z)R(z; \boldsymbol{\theta})} \cdot \frac{1}{U(z)}$$

cannot be computed since it depends on the unknown system G(z) and θ

Idea:

Substitute $R(z; \theta)$ with $R_0(z)$ (that, near the minimum, will be «almost» equal)

$$L(z) = \frac{M(z)}{1 + G(z)R_0(z)} \cdot \frac{W(z)}{U(z)}$$

The asymptotic virtual-reference cost now becomes

$$J_{VR}(\boldsymbol{\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|G|^2 |R_0 - R(\boldsymbol{\theta})|^2}{|1 + GR_0|^2} \cdot \frac{|L|^2}{|M|^2} \cdot \Phi_{uu}(z) \ d\omega$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|G|^2 |R_0 - R(\boldsymbol{\theta})|^2}{|1 + GR_0|^2} \cdot \frac{|M|^2}{|1 + GR_0|^2} \cdot \frac{|W|^2}{\Phi_{uu}} \cdot \Phi_{uu} \cdot \frac{1}{|M|^2} d\omega$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|G|^2}{|1 + GR_0|^2} \cdot \frac{|R_0 - R(\boldsymbol{\theta})|^2}{|1 + GR_0|^2} \cdot |W|^2 d\omega$$

$$J_{\text{MR}}(\boldsymbol{\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|G|^2 |W|^2}{|1 + GR_0|^2} \cdot \frac{|R(\boldsymbol{\theta}) - R_0|^2}{|1 + GR(\boldsymbol{\theta})|^2} d\omega$$

Model-reference cost

$$J_{VR}(\boldsymbol{\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|G|^2 |W|^2}{|1 + GR_0|^2} \cdot \frac{|R_0 - R(\boldsymbol{\theta})|^2}{|1 + GR_0|^2} d\omega$$

Virtual-reference cost

If the family of controllers $R(z; \theta)$ is **only slightly under-parameterized**, it holds that

$$R_0 \approx R(\overline{\boldsymbol{\theta}}), \qquad \overline{\boldsymbol{\theta}} = \operatorname{argmin}_{\boldsymbol{\theta}} J_{\mathrm{MR}}(\boldsymbol{\theta})$$

Then, we can use the approximation

$$1 + G(z)R(z; \boldsymbol{\theta}) \approx 1 + G(z)R_0(z)$$

Since

$$1 - M(z) = \frac{1}{1 + G(z)R_0(z)}$$

we have that

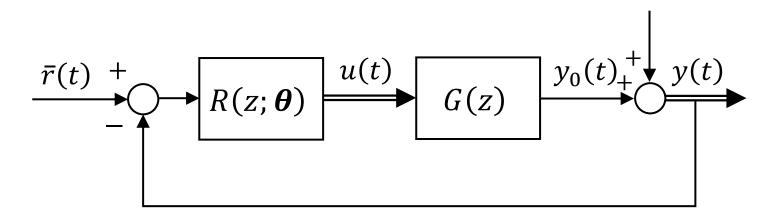
$$L(z) = \frac{M(z)}{1 + G(z)R_0(z)} \cdot \frac{W(z)}{U(z)} = M(z)\left(1 - M(z)\right) \cdot \frac{W(z)}{U(z)}$$

which is a filter perfectly implementable in practice

VRFT for noisy data

In case the plant is affect by **noise** e(t), the output measurements will be

y(t) = G(z)u(t) + e(t)



The regressors

$$\boldsymbol{\varphi}_L^{\mathsf{T}}(t) = \boldsymbol{\beta}^{\mathsf{T}}(z;\boldsymbol{\theta})e_L(t) = \boldsymbol{\beta}^{\mathsf{T}}(z;\boldsymbol{\theta})L(z)e_v(t) = \boldsymbol{\beta}^{\mathsf{T}}(z;\boldsymbol{\theta})L(z)(M^{-1}(z)-1)y(t)$$

are **correlated with the noise** e(t) that acts on y(t). Thus, the **instrumental variable** method has to be employed. The two-stage method can be employed to avoid a second experiment

e(t)

VRFT for noisy data

VRFT summary

- Set $L(z) = (1 M(z))M(z)W(z)U(z)^{-1}$, where $|U(e^{j\omega})|^2 = \Phi_{uu}(\omega)$
- Perform an open-loop experiment on the plant, collecting $\mathcal{D} = \{u(t), y(t)\}_{t=1}^N$
- Compute $u_L(t) = L(z)u(t)$
- Compute $\varphi_L(t) = \beta(z)L(z)(M(z)^{-1} 1)y(t)$
- Identify a high-order model $\widehat{G}(z)$ from \mathcal{D}
- Compute the IV $\mathbf{z}(t) = \boldsymbol{\beta}(z)L(z)(M(z)^{-1} 1)\hat{G}(z)u(t)$
- Estimate the controller parameters $\hat{\theta}_{\text{IV}} = (\sum_{t=1}^{N} \mathbf{z}(t) \boldsymbol{\varphi}^{\mathsf{T}}(t))^{-1} \cdot \sum_{t=1}^{N} \mathbf{z}(t) u_L(t)$

Outline

1. Application study: Virtual Reference Feedback Tuning (VRFT)

2. Defining the control problem

3. Basic idea

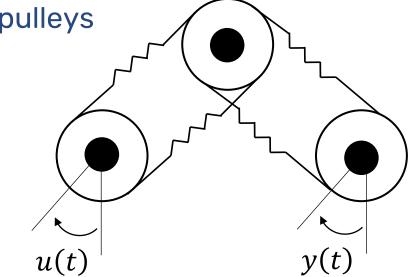
4. VRFT problem solution

5. Example

The flexible transmission consists of three horizontal pulleys connected by two elastic belts

Input: angular position of the first pulley

Output: angular position of the third pulley



The **control objective** is to make the two **angular positions** as **close as possible**. The plant input-output behaviour can be described by (with sampling time $T_s = 0.05 \text{ s}$)

$$G(z) = z^{-3} \cdot \frac{0.28261 + 0.50666z^{-1}}{1 - 1.41833z^{-1} + 1.58939z^{-2} - 1.31608z^{-3} + 0.88642z^{-4}}$$

The control objective is expressed by the following reference model

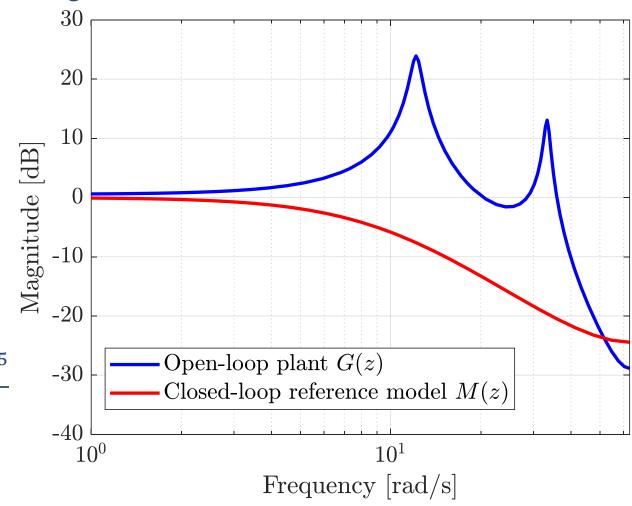
$$M(z) = z^{-3} \cdot \frac{(1-\alpha)^2}{(1-\alpha z^{-1})^2},$$

$$\alpha = e^{-T_S\overline{\omega}}, \quad \overline{\omega} = 10$$

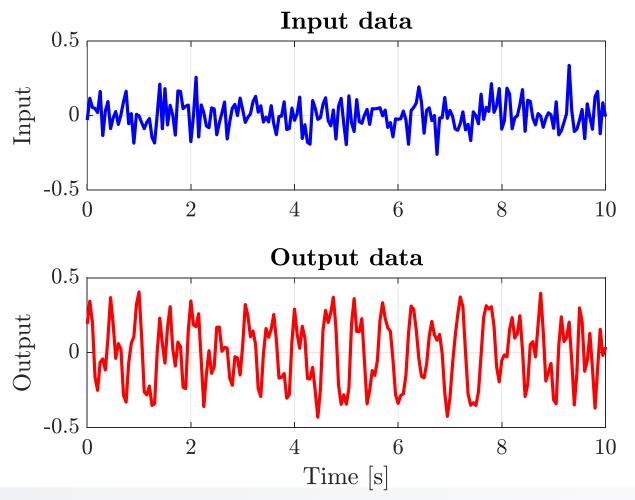
The frequency weight is W(z) = 1

The class of controllers is

$$R(z; \boldsymbol{\theta}) = \frac{\theta_0 + \theta_1 z^{-1} + \theta_2 z^{-2} + \theta_3 z^{-3} + \theta_4 z^{-4} + \theta_5 z^{-5}}{1 - z^{-1}}$$



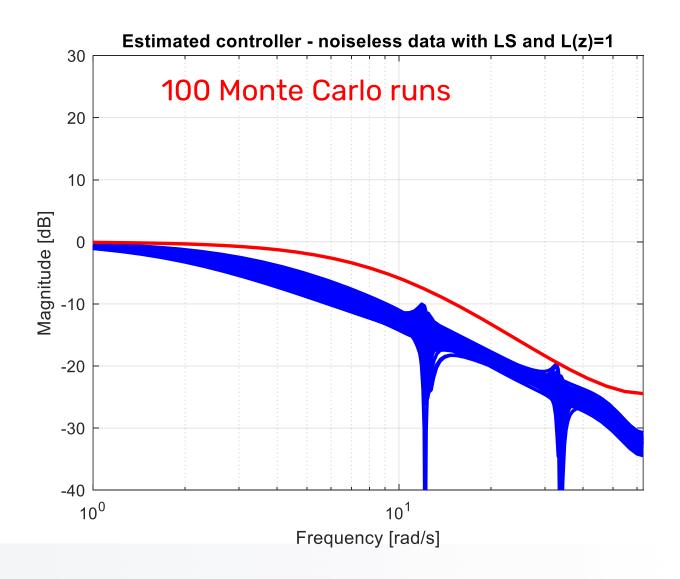
The plant was excited with a white noise input $u(t) \sim WN(0, 0.01)$, collecting N = 512 data



Case 1: no prefilter, noiseless data

$$W(z) = 1, L(z) = 1$$

Not using the optimal prefilter leads to bad results

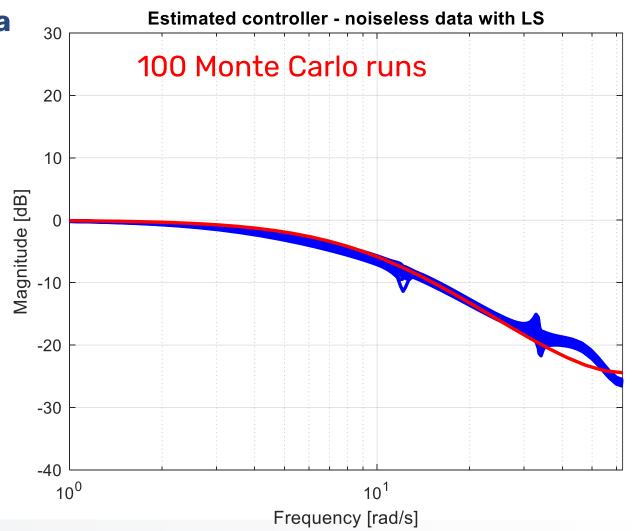


Case 2: optimal prefilter, noiseless data

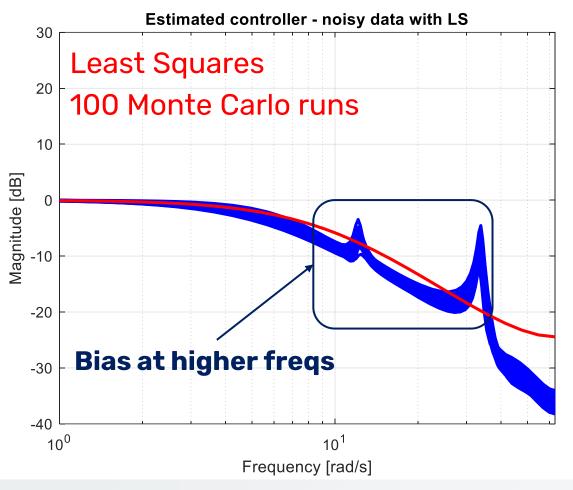
$$W(z) = 1$$
, $\Phi_{uu}(z) = 0.01$

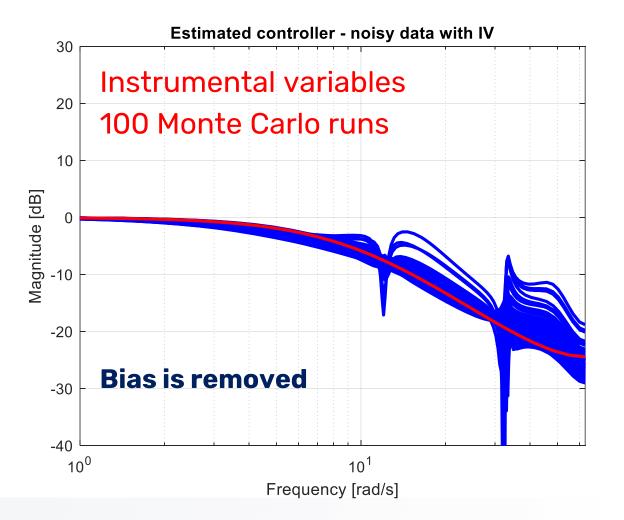
$$L(z) = W(z) \cdot \left(1 - M(z)\right) M(z) \cdot \frac{1}{\Phi_{uu}^{1/2}(z)}$$

The use of the optimal prefilter leads to accurate results



Case 3: optimal prefilter, noisy data







UNIVERSITÀ DEGLI STUDI DI BERGAMO

Dipartimento di Ingegneria Gestionale, dell'Informazione e della Produzione