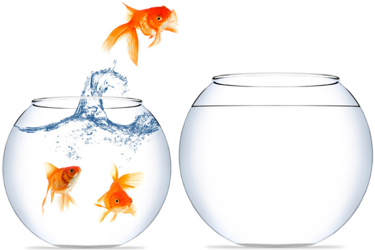
Direct data-driven design of switching controllers

Valentina Breschi

Joint work with S. Formentin @Politecnico di Milano, Italy



ERNSI 2019 WORKSHOP, Maastricht (NL)

September 25, 2019

Outline

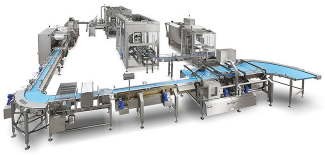
- Problem formulation
- Direct design of switching controllers from data
- A case study
- Concluding remarks

Outline

- Problem formulation
- Direct design of switching controllers from data
- A case study
- Concluding remarks

Introduction

- Many nonlinear and time-varying plants can be described by **switching models**, mixing *discrete* and *continuous dynamics*



- Interpretable and locally simple (e.g. linear) models
- Need for switching identification and control design tools

Piece-wise affine models

- *Input-Output* (IO) description of $\mathcal{G}_s$

$$A(s_{\rm o}(t), q^{-1})y_{\rm o}(t) = B(s_{\rm o}(t), q^{-1})u(t) + f(s_{\rm o}(t))$$

with *switching variable* $s_{\rm o}(t) \in \{1, \dots, K\}$ and

$$A(k, q^{-1}) = 1 + \sum_{i=1}^{n_a} a_i(k)q^{-i} \quad B(k, q^{-1}) = \sum_{i=1}^{n_b} b_i(k)q^{-i}$$

Piece-wise affine models

- *Input-Output* (IO) description of $\mathcal{G}_s$

$$A(s_o(t), q^{-1})y_o(t) = B(s_o(t), q^{-1})u(t) + f(s_o(t))$$

with *switching variable* $s_o(t) \in \{1, \dots, K\}$ and

$$A(k, q^{-1}) = 1 + \sum_{i=1}^{n_a} a_i(k)q^{-i} \quad B(k, q^{-1}) = \sum_{i=1}^{n_b} b_i(k)q^{-i}$$

- Given $X_o(t) = \begin{bmatrix} y_o(t-1) & \cdots & y_o(t-n_a) & u(t-1) & \cdots & u(t-n_b) \end{bmatrix}$ in the domain $\mathcal{X} \subseteq \mathbb{R}^{n_x}$, $\mathcal{G}_s$ changes its operating regime according to a convex polyhedral partition of $\mathcal{X}$, i.e.,

$$s_o(t) = k \iff X_o(t) \in \mathcal{X}_k, \quad \mathcal{X}_k = \{X_o(t), \quad t \in \mathbb{N} : \mathcal{H}_k X_o(t) \preceq \mathcal{D}_k\}$$

Piece-wise affine models

- *Identification*: known to be **NP-hard**. Several heuristic (usually 2-step) procedures have been proposed.
- *Control*: several contributions, among which Lyapunov-based, LMI-based and Hybrid MPC design
- *Fixed-order control*: few works on switching PID tuning

Open problems

- Model reduction for piece-wise affine models is hard (especially in case of endogenous switching)
- The effect of modeling errors on the closed-loop performance is *unknown*
- The state of the system is assumed to be available, but this might not be true in practice

In this talk

- Direct design of (model-reference) fixed-order switching controllers from data

  ✓ no assumptions on the model structure
  ✓ control-oriented cost optimization
  ✓ no model/controller reduction needed

  ✗ controller parameters/partition are no longer linked to the physics

Assumptions

A1 A set of IO data $\mathcal{D}_T = \{u(t), y(t)\}_{t=1}^{T}$ is available

A2 $\mathcal{G}_s$ is BIBO stable

A3 The desired closed-loop behavior $\mathcal{M}_s$ is given

$$
\begin{aligned}
x_M(t+1) &= A_M(s_M(t))x_M(t) + B_M(s_M(t))r(t) \\
y_d(t) &= C_M(s_M(t))x_M(t) + D_M(s_M(t))r(t)
\end{aligned}
$$

$$
s_M(t) = i \iff x_M(t) \in \mathcal{X}_i^M, \quad i \in \{1, \dots, K_M\}.
$$

Problem statement

Problem (*Data-driven design of switching controllers*)

Given a set of dataset $\mathcal{D}_T = \{u(t), y(t)\}_{t=1}^{T}$ *collected from open-loop experiments on* $\mathcal{G}_s$ *and a reference closed-loop model* $\mathcal{M}_s$*, find the PWA controller with fixed-structure*

$$A_c(s_c(t), q^{-1})u(t) = B_c(s_c(t), q^{-1})e(t) + f_c(s_c(t)),$$

where $e(t) = r(t) - y(t)$*,* $s_c(t) \in \{1, \ldots, K_c\}$*,* $K_c \in \mathbb{N}$*,*

$$A_c(k, q^{-1}) = 1 + \sum_{i=1}^{n_a^c} a_i^c(k)q^{-i}, \quad B_c(k, q^{-1}) = \sum_{i=0}^{n_b^c} b_i^c(k)q^{-i},$$

$$\chi(t) = \begin{bmatrix} u(t-1) & \cdots & u(t-n_a^c) & y(t) & y(t-1) & \ldots y(t-n_b^c) \end{bmatrix}' \in \mathcal{X}^c \subseteq \mathbb{R}^{n_\chi},$$

$$s_c(t) = j \iff \chi(t) \in \mathcal{X}_j^c$$

that realizes the desired closed-loop behavior $\mathcal{M}_s$*.*

Problem statement(cont'd)

$$A_c(s_c(t), q^{-1})u(t) = B_c(s_c(t), q^{-1})e(t) + f_c(s_c(t)),$$

$$s_c(t) = j \iff \chi(t) \in \mathcal{X}_j^c$$

Unknowns:

- Number of local controllers K_c

Problem statement(cont'd)

$$A_c(s_c(t), q^{-1})u(t) = B_c(s_c(t), q^{-1})e(t) + f_c(s_c(t)),$$

$$s_c(t) = j \iff \chi(t) \in \mathcal{X}_j^c$$

Unknowns:

- Number of local controllers K_c
- Parameters of the local controllers

Problem statement(cont'd)

$$A_c(s_c(t), q^{-1})u(t) = B_c(s_c(t), q^{-1})e(t) + f_c(s_c(t)),$$

$$s_c(t) = j \iff \chi(t) \in \mathcal{X}_j^c$$

Unknowns:

- Number of local controllers K_c
- Parameters of the local controllers
- Partition of the regressor space $\mathcal{X}^c$

Illustrative example

$$x(t+1) = 0.8 \begin{bmatrix} \cos\alpha(t) & -\sin\alpha(t) \\ \sin\alpha(t) & \cos\alpha(t) \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t),$$

$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(t),$$

where

$$\alpha(t) = \begin{cases} \frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) \geq 0, \\ -\frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) < 0. \end{cases}$$

- The system is **piece-wise linear**
- The reference model is **LTI**

$$x_M(t+1) = \begin{bmatrix} -1 & 1 \\ -0.8 & 1 \end{bmatrix} x_M(t) + \begin{bmatrix} 0.6 \\ 0.8 \end{bmatrix} r(t),$$

$$y_d(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x_M(t).$$

Illustrative example

$$x(t+1) = 0.8 \begin{bmatrix} \cos\alpha(t) & -\sin\alpha(t) \\ \sin\alpha(t) & \cos\alpha(t) \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t),$$
$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(t),$$

where

$$\alpha(t) = \begin{cases} \frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) \geq 0, \\ -\frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) < 0. \end{cases}$$

- The system is **piece-wise linear**
- The reference model is **LTI**

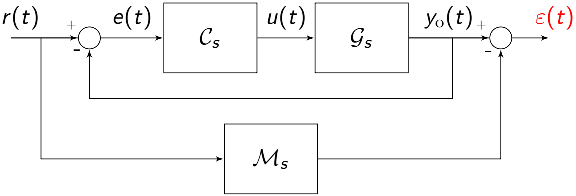
Selected fixed-order controller structure

$$u(t) = u(t-1) + \theta_{s_c(t),1} \overbrace{e(t)}^{r(t)-y(t)} + \theta_{s_c(t),2} e(t-1) + \theta_{s_c(t),3},$$
$$s_c(t) = i \iff \chi(t) = \begin{bmatrix} u(t-1) & y(t) & y(t-1) \end{bmatrix}' \in \mathcal{X}_i^c \text{ for } i = 1, 2$$

Outline

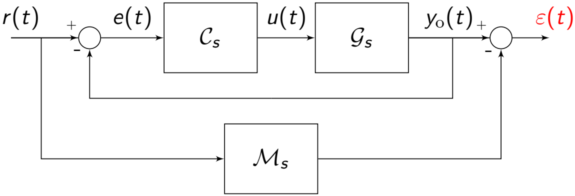
- Problem formulation
- Direct design of switching controllers from data
- A case study
- Concluding remarks

Direct design of switching controllers from data



- **Goal**: design $\mathcal{C}_s$ to minimize the *mismatch error* ε

Direct design of switching controllers from data



$$\min_{\varepsilon,\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \|\varepsilon\|_{\ell_2}^2$$

$$\begin{aligned}
s.t.\ \ & \varepsilon(t) = y_o(t) - M(s_M(t))r(t),\ \forall t \in \mathcal{I}_1^T \\
& A(s_o(t))y_o(t) = B(s_o(t))u(t) + f(s_o(t)),\ \forall t \in \mathcal{I}_1^T \\
& s_o(t) = k \iff X_o(t) \in \mathcal{X}_k, k \in \{1,\dots,K\},\ \forall t \in \mathcal{I}_1^T \\
& A_c(s_c(t))u(t) = B_c(s_c(t))(r(t) - y_o(t)) + f_c(s_c(t)),\ \forall t \in \mathcal{I}_1^T \\
& s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c, i \in \{1,\dots,K_c\},\ \forall t \in \mathcal{I}_1^T
\end{aligned}$$

Removing the dependence on the reference

$$\begin{aligned}
\min_{\varepsilon,\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} & \|\varepsilon\|_{\ell_2}^2 \\
s.t. \;\; & \varepsilon(t) = y_o(t) - M(s_M(t))r(t), \;\; \forall t \in \mathcal{I}_1^T \\
& A(s_o(t))y_o(t) = B(s_o(t))u(t) + f(s_o(t)), \;\; \forall t \in \mathcal{I}_1^T \\
& s_o(t) = k \iff X_o(t) \in \mathcal{X}_k, k \in \{1,\ldots,K\}, \;\; \forall t \in \mathcal{I}_1^T \\
& A_c(s_c(t))u(t) = B_c(s_c(t))(r(t) - y_o(t)) + f_c(s_c(t)), \;\; \forall t \in \mathcal{I}_1^T \\
& s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c, i \in \{1,\ldots,K_c\}, \;\; \forall t \in \mathcal{I}_1^T
\end{aligned}$$

- The *explicit* dependence on r implies that the problem needs to be solved again every time a different signal is considered.

Removing the dependence on the reference

$$\min_{\varepsilon,\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \|\varepsilon\|_{\ell_2}^2$$

$$\begin{aligned}
s.t.\ \ & \varepsilon(t) = y_o(t) - M(s_M(t))r(t),\ \forall t \in \mathcal{I}_1^T \\
& A(s_o(t))y_o(t) = B(s_o(t))u(t) + f(s_o(t)),\ \forall t \in \mathcal{I}_1^T \\
& s_o(t) = k \iff X_o(t) \in \mathcal{X}_k, k \in \{1,\dots,K\},\ \forall t \in \mathcal{I}_1^T \\
& A_c(s_c(t))u(t) = B_c(s_c(t))(r(t) - y_o(t)) + f_c(s_c(t)),\ \forall t \in \mathcal{I}_1^T \\
& s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c, i \in \{1,\dots,K_c\},\ \forall t \in \mathcal{I}_1^T
\end{aligned}$$

- *Virtual reference* $\tilde{r}(t) = M^\dagger(s_M(t))y_o(t) - M^\dagger(s_M(t))\varepsilon(t)$
- $\tilde{e}(t) = \tilde{r}(t) - y_o(t) = (M^\dagger(s_M(t)) - 1)y_o(t) - M^\dagger(s_M(t))\varepsilon(t)$

➜ The reference depends on the user defined $M^\dagger(s_M(t))$, the unknown $\varepsilon(t)$ and the noiseless output $y_o(t)$.

Removing the dependence on the reference

- *Virtual reference* $\tilde{r}(t) = M^{\dagger}(s_M(t))y_{\text{o}}(t) - M^{\dagger}(s_M(t))\varepsilon(t)$
- $\tilde{e}(t) = \tilde{r}(t) - y_{\text{o}}(t) = (M^{\dagger}(s_M(t)) - 1)y_{\text{o}}(t) - M^{\dagger}(s_M(t))\varepsilon(t)$

➜
$$\min_{\varepsilon,\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \|\varepsilon\|_{\ell_2}^2$$

$$\begin{aligned}
s.t.\ \ & A(s_{\text{o}}(t))y_{\text{o}}(t) = B(s_{\text{o}}(t))u(t) + f(s_{\text{o}}(t)),\ \forall t \in \mathcal{I}_1^T \\
& s_{\text{o}}(t) = k \iff X_{\text{o}}(t) \in \mathcal{X}_k, k \in \{1, \dots, K\},\ \forall t \in \mathcal{I}_1^T, \\
& A_c(s_c(t))u(t) = B_c(s_c(t))\tilde{e}(t) + f_c(s_c(t)),\ \forall t \in \mathcal{I}_1^T \\
& s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c,\ i \in \{1, \dots, K_c\},\ \forall t \in \mathcal{I}_1^T,
\end{aligned}$$

➜ The problem can be solved **independently** of the reference

Computation of $M^\dagger$

$$\min_{\varepsilon,\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \|\varepsilon\|_{\ell_2}^2$$

$$\begin{aligned}
s.t.\;\; & A(s_\mathrm{o}(t))y_\mathrm{o}(t) = B(s_\mathrm{o}(t))u(t) + f(s_\mathrm{o}(t)),\ \forall t \in \mathcal{I}_1^T \\
& s_\mathrm{o}(t) = k \iff X_\mathrm{o}(t) \in \mathcal{X}_k, k \in \{1,\ldots,K\},\ \forall t \in \mathcal{I}_1^T, \\
& A_c(s_c(t))u(t) = B_c(s_c(t))\tilde{e}(t) + f_c(s_c(t)),\ \forall t \in \mathcal{I}_1^T \\
& s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c,\ i \in \{1,\ldots,K_c\},\ \forall t \in \mathcal{I}_1^T,
\end{aligned}$$

- The problem depends on the computation of $M^\dagger$

Computation of $M^\dagger$

$$\min_{\varepsilon,\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \|\varepsilon\|_{\ell_2}^2$$

$$\begin{aligned}
s.t.\ \ & A(s_\mathrm{o}(t))y_\mathrm{o}(t) = B(s_\mathrm{o}(t))u(t) + f(s_\mathrm{o}(t)),\ \forall t \in \mathcal{I}_1^T \\
& s_\mathrm{o}(t) = k \iff X_\mathrm{o}(t) \in \mathcal{X}_k, k \in \{1,\ldots,K\},\ \forall t \in \mathcal{I}_1^T, \\
& A_c(s_c(t))u(t) = B_c(s_c(t))\tilde{e}(t) + f_c(s_c(t)),\ \forall t \in \mathcal{I}_1^T \\
& s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c,\ i \in \{1,\ldots,K_c\},\ \forall t \in \mathcal{I}_1^T,
\end{aligned}$$

- The problem depends on the computation of $M^\dagger$

➜ $\mathcal{M}_s$ is an LPV model with integer scheduling variable

The left-inverse can be computed similarly to the LPV case

(Formentin, Piga, Tóth, Savaresi, 2016)

Removing the dependence on the model of $\mathcal{G}_s$

$$\min_{\varepsilon,\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \|\varepsilon\|_{\ell_2}^2$$

$$\begin{aligned} s.t.\ & A(s_\mathrm{o}(t))y_\mathrm{o}(t) = B(s_\mathrm{o}(t))u(t) + f(s_\mathrm{o}(t)),\ \forall t \in \mathcal{I}_1^T \\ & s_\mathrm{o}(t) = k \iff X_\mathrm{o}(t) \in \mathcal{X}_k,\ k \in \{1,\dots,K\},\ \forall t \in \mathcal{I}_1^T, \\ & A_c(s_c(t))u(t) = B_c(s_c(t))\tilde{e}(t) + f_c(s_c(t)),\ \forall t \in \mathcal{I}_1^T \\ & s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c,\ i \in \{1,\dots,K_c\},\ \forall t \in \mathcal{I}_1^T, \end{aligned}$$

- The problem is model-based

Removing the dependence on the model of $\mathcal{G}_s$

$$\min_{\varepsilon,\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \|\varepsilon\|_{\ell_2}^2$$

$$\begin{aligned}
s.t.\ \ & A(s_{\rm o}(t))y_{\rm o}(t) = B(s_{\rm o}(t))u(t) + f(s_{\rm o}(t)),\ \forall t \in \mathcal{I}_1^T \\
& s_{\rm o}(t) = k \iff X_{\rm o}(t) \in \mathcal{X}_k,\ k \in \{1,\dots,K\},\ \forall t \in \mathcal{I}_1^T, \\
& A_c(s_c(t))u(t) = B_c(s_c(t))\tilde{e}(t) + f_c(s_c(t)),\ \forall t \in \mathcal{I}_1^T \\
& s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c,\ i \in \{1,\dots,K_c\},\ \forall t \in \mathcal{I}_1^T,
\end{aligned}$$

- Leverage on the available dataset $\mathcal{D}_{\mathbf{T}} = \{\mathbf{u(t)},\mathbf{y(t)}\}_{\mathbf{t=1}}^{\mathbf{T}}$
- $\tilde{r}(t) = M^{\dagger}(s_M(t))y(t) - M^{\dagger}(s_M(t))\varepsilon(t)$
- $\tilde{e}(t) = \tilde{r}(t) - y(t) = (M^{\dagger}(s_M(t)) - 1)y(t) - M^{\dagger}(s_M(t))\varepsilon(t)$

Removing the dependence on the model of $\mathcal{G}_s$

- Leverage on the available dataset $\mathcal{D}_\mathbf{T} = \{\mathbf{u}(\mathbf{t}), \mathbf{y}(\mathbf{t})\}_{\mathbf{t}=\mathbf{1}}^{\mathbf{T}}$
- $\tilde{r}(t) = M^\dagger(s_M(t))y(t) - M^\dagger(s_M(t))\varepsilon(t)$
- $\tilde{e}(t) = \tilde{r}(t) - y(t) = (M^\dagger(s_M(t)) - 1)y(t) - M^\dagger(s_M(t))\varepsilon(t)$

$$
\begin{aligned}
\min_{\varepsilon,\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \quad & \|\varepsilon\|_{\ell_2}^2 \\
s.t. \quad & A_c(s_c(t))u(t) = B_c(s_c(t))\tilde{e}(t) + f_c(s_c(t)), \\
& s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c, i \in \{1, \dots, K_c\}, \{u(t), y(t)\} \in \mathcal{D}_T.
\end{aligned}
$$

➜ The problem depends on the user defined $M^\dagger(s_M(t))$ and the unknown $\varepsilon(t)$.

Removing the dependence on the model of $\mathcal{G}_s$ (cont'd)

$$\min_{\varepsilon,\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \|\varepsilon\|_{\ell_2}^2$$
$$s.t.\ \ A_c(s_c(t))u(t) = B_c(s_c(t))\tilde{e}(t) + f_c(s_c(t)),$$
$$s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c, i \in \{1,\ldots,K_c\}, \{u(t),y(t)\} \in \mathcal{D}_T.$$

- The problem is bi-convex

Removing the dependence on the model of $\mathcal{G}_s$ (cont'd)

$$\min_{\varepsilon,\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \|\varepsilon\|_{\ell_2}^2$$
$$\text{s.t. } A_c(s_c(t))u(t) = B_c(s_c(t))\tilde{e}(t) + f_c(s_c(t)),$$
$$s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c, i \in \{1,\ldots,K_c\}, \{u(t),y(t)\} \in \mathcal{D}_T.$$

$$B_c(s_c(t))M^\dagger\varepsilon(t) = B_c(s_c(t))\underbrace{(M^\dagger - 1)y(t)}_{\xi(t)} - A_c(s_c(t))u(t) + f_c(s_c(t)),$$

$$\Downarrow$$

$$B_c(s_c(t),q^{-1})M^\dagger(s_M(t))\varepsilon(t) = \underbrace{X_c(\xi(t))'}_{\left[-u(t-1)\ \cdots\ -u(t-n_a^c)\ \ \xi(t)\ \cdots\ \xi(t-n_b^c)\ \ 1\right]} \theta_{s_c(t)} - u(t).$$

Removing the dependence on the model of $\mathcal{G}_s$ (cont'd)

$$B_c(s_c(t), q^{-1})M^{\dagger}(s_M(t))\varepsilon(t) = X_c(\xi(t))'\theta_{s_c(t)} - u(t).$$

$$\min_{\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \sum_{t=1}^{T} \left(X_c(\xi(t))'\theta_{s_c(t)} - u(t)\right)^2 + \lambda\|\Theta\|_2^2$$

$$s.t.\ \ s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c, i \in \{1,\ldots,K_c\}, \{u(t), y(t)\} \in \mathcal{D}_T.$$

To avoid the trivial solution $B_c(k) = 0$, for all $k = 1, \ldots, K$, parameter estimation is performed via *instrumental variables*

Explicit estimation of the polyhedral partition of $\mathcal{X}^c$

$$\min_{\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \sum_{t=1}^{T} \left(X_c(\xi(t))'\theta_{s_c(t)} - u(t)\right)^2 + \lambda\|\Theta\|_2^2$$

$$s.t.\;\; s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c, i \in \{1,\ldots,K_c\}, \{u(t), y(t)\} \in \mathcal{D}_T.$$

- The problem is constrained by the (unknown) switching logic of the controller.

➜ Introduce in the cost the explicit dependence on the (unknown) polyhedral partition of $\mathcal{X}^c$

Explicit estimation of the polyhedral partition of $\mathcal{X}^c$

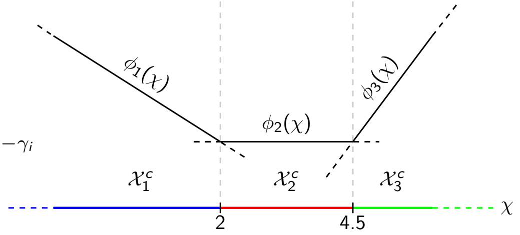
$$\min_{\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \sum_{t=1}^{T} \left(X_c(\xi(t))'\theta_{s_c(t)} - u(t)\right)^2 + \lambda\|\Theta\|_2^2$$

$$s.t. \;\; s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c, i \in \{1,\ldots,K_c\}, \{u(t),y(t)\} \in \mathcal{D}_T.$$

- The polyhedral partition of $\mathcal{X}^c$ can be estimated by solving a **multicategory discrimination** problem

➜ **Goal:** Find a PWA separator

$$\phi(\chi) = \max_{i=1,\ldots,K_c} \underbrace{\phi_i(\chi)}_{\phi_i(\chi)=(\omega_i)'\chi-\gamma_i}$$



Explicit estimation of the polyhedral partition of $\mathcal{X}^c$

$$\min_{\Theta,\{\mathcal{X}_i^c\}_{i=1}^{K_c},\mathcal{S}_c} \sum_{t=1}^{T} \left(X_c(\xi(t))'\theta_{s_c(t)} - u(t)\right)^2 + \lambda\|\Theta\|_2^2$$

$$\text{s.t.} \quad s_c(t) = i \iff \chi(t) \in \mathcal{X}_i^c, i \in \{1,\ldots,K_c\}, \{u(t),y(t)\} \in \mathcal{D}_T.$$

$\Downarrow$

$$\min_{\Theta,\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c} \mathcal{J}_1(\Theta,\mathcal{S}_c) + \mathcal{J}_2(\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c).$$

$$\mathcal{J}_1(\Theta,\mathcal{S}_c) = \sum_{t=1}^{T} \left(X_c(\xi(t))'\theta_{s_c(t)} - u(t)\right)^2 + \lambda\|\Theta\|_2^2$$

$$\mathcal{J}_2(\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c) = \sum_{t=1}^{T}\sum_{\substack{j=1\\ j\neq s_c(t)}}^{K_c} \left\| \left(\begin{bmatrix}\chi(t)' & -1\end{bmatrix} \begin{bmatrix}\omega_j - \omega_{s_c(t)} \\ \gamma_j - \gamma_{s_c(t)}\end{bmatrix} + 1 \right)_+ \right\|_2^2 + \lambda_p \sum_{i=1}^{K_c} (\|\omega_i\|_2^2 + \gamma_i^2)$$

Data-driven design of PWA controllers

$$\min_{\Theta,\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c} \mathcal{J}_1(\Theta,\mathcal{S}_c)+\mathcal{J}_2(\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c)$$

$$\mathcal{J}_1(\Theta,\mathcal{S}_c)=\sum_{t=1}^{T}\left(X_c(\xi(t))'\theta_{s_c(t)}-u(t)\right)^2+\lambda\|\Theta\|_2^2$$

$$\mathcal{J}_2(\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c)=\sum_{t=1}^{T}\sum_{\substack{j=1\\ j\neq s_c(t)}}^{K_c}\left\|\left(\left[\chi(t)'\ -1\right]\begin{bmatrix}\omega_j-\omega_{s_c(t)}\\ \gamma_j-\gamma_{s_c(t)}\end{bmatrix}+1\right)_+\right\|_2^2+\lambda_p\sum_{i=1}^{K_c}(\|\omega_i\|_2^2+\gamma_i^2)$$

Data-driven design of PWA controllers

$$\min_{\Theta,\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c} \mathcal{J}_1(\Theta,\mathcal{S}_c) + \mathcal{J}_2(\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c)$$

$$\mathcal{J}_1(\Theta,\mathcal{S}_c) = \sum_{t=1}^{T}\left(X_c(\xi(t))'\theta_{s_c(t)} - u(t)\right)^2 + \lambda\|\Theta\|_2^2$$

$$\mathcal{J}_2(\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c) = \sum_{t=1}^{T}\sum_{\substack{j=1\\ j\neq s_c(t)}}^{K_c}\left\|\left(\left[\chi(t)' \; -1\right]\begin{bmatrix}\omega_j-\omega_{s_c(t)}\\ \gamma_j-\gamma_{s_c(t)}\end{bmatrix}+1\right)_+\right\|_2^2 + \lambda_p\sum_{i=1}^{K_c}(\|\omega_i\|_2^2+\gamma_i^2)$$

- **Step 1**: estimation of the parameters Θ of the local controllers (supervised for fixed $\mathcal{S}_c$)

Data-driven design of PWA controllers

$$\min_{\Theta,\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c} \mathcal{J}_1(\Theta,\mathcal{S}_c) + \mathcal{J}_2(\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c)$$

$$\mathcal{J}_1(\Theta,\mathcal{S}_c) = \sum_{t=1}^{T}\left(X_c(\xi(t))'\theta_{s_c(t)} - u(t)\right)^2 + \lambda\|\Theta\|_2^2$$

$$\mathcal{J}_2(\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c)=\sum_{t=1}^{T}\sum_{\substack{j=1\\ j\neq s_c(t)}}^{K_c}\left\|\left(\left[\chi(t)'\ -1\right]\begin{bmatrix}\omega_j-\omega_{s_c(t)}\\ \gamma_j-\gamma_{s_c(t)}\end{bmatrix}+1\right)_+\right\|_2^2+\lambda_p\sum_{i=1}^{K_c}(\|\omega_i\|_2^2+\gamma_i^2)$$

- **Step 1**: estimation of the parameters Θ of the local controllers (supervised for fixed $\mathcal{S}_c$)
- **Step 2**: estimation of the parameters $\{\omega_i,\gamma_i\}_{i=1}^{K_c}$ of the PWA separator (supervised for fixed $\mathcal{S}_c$)

Data-driven design of PWA controllers

$$\min_{\Theta,\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c} \mathcal{J}_1(\Theta,\mathcal{S}_c)+\mathcal{J}_2(\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c)$$

$$\mathcal{J}_1(\Theta,\mathcal{S}_c)=\sum_{t=1}^{T}\left(X_c(\xi(t))'\theta_{s_c(t)}-u(t)\right)^2+\lambda\|\Theta\|_2^2$$

$$\mathcal{J}_2(\{\omega_i,\gamma_i\}_{i=1}^{K_c},\mathcal{S}_c)=\sum_{t=1}^{T}\sum_{\substack{j=1\\ j\neq s_c(t)}}^{K_c}\left\|\left(\begin{bmatrix}\chi(t)' & -1\end{bmatrix}\begin{bmatrix}\omega_j-\omega_{s_c(t)}\\ \gamma_j-\gamma_{s_c(t)}\end{bmatrix}+1\right)_+\right\|_2^2+\lambda_p\sum_{i=1}^{K_c}(\|\omega_i\|_2^2+\gamma_i^2)$$

- **Step 1**: estimation of the parameters Θ of the local controllers (supervised for fixed $\mathcal{S}_c$)
- **Step 2**: estimation of the parameters $\{\omega_i,\gamma_i\}_{i=1}^{K_c}$ of the PWA separator (supervised for fixed $\mathcal{S}_c$)
- **Step 3**: update the mode sequence $\mathcal{S}_c$ for fixed Θ and $\{\omega_i,\gamma_i\}_{i=1}^{K_c}$ (unsupervised)

Step 1: Estimation of the parameters Θ

For $\mathcal{S}_c = \mathcal{S}_c^{k-1}$:

$$\min_{\Theta} \sum_{t=1}^{T} \left(X_c(\xi(t))' \theta_{s_c^{k-1}(t)} - u(t) \right)^2 + \lambda \|\Theta\|_2^2$$

✓ The problem is **separable**

✗ The solution is **biased** even when the *ideal* mode sequence is known

✗ The solution might correspond to the **trivial** one ($B_c(k) = 0$)

⬇

Instrumental variable scheme

➤ Instrument $\zeta(t)$ s.t. $\mathbb{E}[\zeta(t)(\xi - \xi_o)] = 0$, $\forall t$

$$\min_{\Theta} \left\| \sum_{t=1}^{T} \zeta(t) \left(X_c(\xi(t))' \theta_{s_c^{k-1}(t)} - u(t) \right) \right\|_2^2 + \lambda \|\Theta\|_2^2.$$

Step 1: Estimation of the parameters Θ (cont'd)

$$\min_{\Theta} \left\| \sum_{t=1}^{T} \zeta(t) \left(X_c(\xi(t))' \theta_{s_c^{k-1}(t)} - u(t) \right) \right\|_2^2 + \lambda \|\Theta\|_2^2.$$

✗ The problem is **not separable**

- ➤ Batch solution

⬇

$$\min_{\Theta} \left\| \sum_{t=1}^{T} \zeta(t) \left(X_c(\xi(t))' \mathbb{I}_{[s_c^{k-1}(t)]} \Theta - u(t) \right) \right\|_2^2 + \lambda \|\Theta\|_2^2,$$

Step 1: Estimation of the parameters Θ (cont'd)

$$\min_{\Theta} \left\| \sum_{t=1}^{T} \zeta(t) \left(X_c(\xi(t))' \mathbb{I}_{[s_c^{k-1}(t)]} \Theta - u(t) \right) \right\|_2^2 + \lambda \|\Theta\|_2^2,$$

⬇

$$\min_{\Theta} \|Z'(R(\xi)\Theta - U)\|_2^2 + \lambda \|\Theta\|_2^2,$$

➜ $\Theta = \left((Z'R(\xi))'Z'R(\xi) + \lambda I_{n_\Theta}\right)^{-1} (Z'R(\xi))'Z'U$

Step 1: Estimation of the parameters Θ (cont'd)

$$\min_{\Theta} \left\| \sum_{t=1}^{T} \zeta(t) \left(X_c(\xi(t))' \mathbb{I}_{[s_c^{k-1}(t)]} \Theta - u(t) \right) \right\|_2^2 + \lambda \|\Theta\|_2^2,$$

$$\min_{\Theta} \|Z'(R(\xi)\Theta - U)\|_2^2 + \lambda \|\Theta\|_2^2,$$

➜ $\Theta = \left((Z'R(\xi))'Z'R(\xi) + \lambda I_{n_\Theta}\right)^{-1} (Z'R(\xi))'Z'U$

If

- The instrument is properly chosen
- The number of **misclassified points** is small $(|\{t : \mathcal{S}_c \neq \mathcal{S}_c^0\}| \ll T)$

Proposition

$$\lim_{T \to \infty} \hat{\Theta} = \Theta^{\mathrm{o}}$$

Step 1: illustrative example

$$x(t+1) = 0.8 \begin{bmatrix} \cos\alpha(t) & -\sin\alpha(t) \\ \sin\alpha(t) & \cos\alpha(t) \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t),$$
$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(t),$$

where

$$\alpha(t) = \begin{cases} \frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) \geq 0, \\ -\frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) < 0. \end{cases}$$

Chosen fixed-order controller structure

$$u(t) = u(t-1) + \theta_{s_c(t),1} e(t) + \theta_{s_c(t),2} e(t-1) + \theta_{s_c(t),3},$$
$$s_c(t) = i \iff \chi(t) = \begin{bmatrix} u(t-1) & y(t) & y(t-1) \end{bmatrix}' \in \mathcal{X}_i^c \text{ for } i = 1, 2$$

Step 1: illustrative example

$$x(t+1) = 0.8 \begin{bmatrix} \cos\alpha(t) & -\sin\alpha(t) \\ \sin\alpha(t) & \cos\alpha(t) \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t),$$

$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(t),$$

where

$$\alpha(t) = \begin{cases} \frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) \geq 0, \\ -\frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) < 0. \end{cases}$$

Chosen fixed-order controller structure

$$u(t) = u(t-1) + \theta_{s_c(t),1} e(t) + \theta_{s_c(t),2} e(t-1) + \theta_{s_c(t),3},$$

$$s_c(t) = i \iff \chi(t) = \begin{bmatrix} u(t-1) & y(t) & y(t-1) \end{bmatrix}' \in \mathcal{X}_i^c \text{ for } i = 1, 2$$

- **Estimate** the parameters of the local controllers, with

$$X_c(\xi(t)) = \begin{bmatrix} -u(t-1) & (M^\dagger(s_M(t))-1)y(t) & (M^\dagger(s_M(t-1))-1)y(t-1) & 1 \end{bmatrix}'$$

- **Fix** the partition $\{\mathcal{X}_i^c\}_{i=1}^{K_c}$ and the mode sequence $\mathcal{S}_c = \{s_c(t)\}_{t=1}^{T}$

Step 2: learning the PWA separator

- For **fixed** $\mathcal{S}_c = \mathcal{S}_c^{k-1}$, we can construct K_c **clusters** $\{C_i\}_{i=1}^{K_c}$

➜ The **clusters** are used to **estimate** the PWA separator

$$\min_{\{\omega_i,\gamma_i\}_{i=1}^{K_c}} \sum_{i=1}^{K_c} \sum_{C_i} \sum_{\substack{j=1 \\ j\neq i}}^{K_c} \left\| \left(\begin{bmatrix} \chi(t)' & -1 \end{bmatrix} \begin{bmatrix} \omega_j - \omega_i \\ \gamma_j - \gamma_i \end{bmatrix} + 1 \right)_+ \right\|_2^2 + \lambda_p \sum_{i=1}^{K_c} (\|\omega_i\|_2^2 + \gamma_i^2),$$

Step 2: learning the PWA separator

- For **fixed** $\mathcal{S}_c = \mathcal{S}_c^{k-1}$, we can construct K_c **clusters** $\{C_i\}_{i=1}^{K_c}$

➜ The **clusters** are used to **estimate** the PWA separator

$$\min_{\{\omega_i,\gamma_i\}_{i=1}^{K_c}} \sum_{i=1}^{K_c} \sum_{C_i} \sum_{\substack{j=1 \\ j\neq i}}^{K_c} \left\| \left(\begin{bmatrix} \chi(t)' & -1 \end{bmatrix} \begin{bmatrix} \omega_j - \omega_i \\ \gamma_j - \gamma_i \end{bmatrix} + 1 \right)_+ \right\|_2^2 + \lambda_p \sum_{i=1}^{K_c} (\|\omega_i\|_2^2 + \gamma_i^2),$$

⬇

$$\min_{\{\omega_i,\gamma_i\}_{i=1}^{K_c}} \sum_{i=1}^{K_c} \sum_{\substack{j=1 \\ j\neq i}}^{K_c} \frac{1}{c_i} \left\| \left(\begin{bmatrix} X_i' & -\mathbf{1}_{c_i} \end{bmatrix} \begin{bmatrix} \omega_j - \omega_i \\ \gamma_j - \gamma_i \end{bmatrix} + \mathbf{1}_{c_i} \right)_+ \right\|_2^2 + \lambda_p \sum_{i=1}^{K_c} (\|\omega_i\|_2^2 + \gamma_i^2),$$

Step 2: learning the PWA separator

- For **fixed** $\mathcal{S}_c = \mathcal{S}_c^{k-1}$, we can construct K_c **clusters** $\{C_i\}_{i=1}^{K_c}$

➜ The **clusters** are used to **estimate** the PWA separator

$$\min_{\{\omega_i,\gamma_i\}_{i=1}^{K_c}} \sum_{i=1}^{K_c} \sum_{C_i} \sum_{\substack{j=1 \\ j\neq i}}^{K_c} \left\| \left(\begin{bmatrix} \chi(t)' & -1 \end{bmatrix} \begin{bmatrix} \omega_j - \omega_i \\ \gamma_j - \gamma_i \end{bmatrix} + 1 \right)_+ \right\|_2^2 + \lambda_p \sum_{i=1}^{K_c} (\|\omega_i\|_2^2 + \gamma_i^2),$$

⬇

$$\min_{\{\omega_i,\gamma_i\}_{i=1}^{K_c}} \sum_{i=1}^{K_c} \sum_{\substack{j=1 \\ j\neq i}}^{K_c} \frac{1}{c_i} \left\| \left(\begin{bmatrix} X_i' & -\mathbf{1}_{c_i} \end{bmatrix} \begin{bmatrix} \omega_j - \omega_i \\ \gamma_j - \gamma_i \end{bmatrix} + \mathbf{1}_{c_i} \right)_+ \right\|_2^2 + \lambda_p \sum_{i=1}^{K_c} (\|\omega_i\|_2^2 + \gamma_i^2),$$

- X_i is constructed according to the cluster C_i, $i = 1, \ldots, K_c$

Step 2: learning the PWA separator

- For **fixed** $\mathcal{S}_c = \mathcal{S}_c^{k-1}$, we can construct K_c **clusters** $\{C_i\}_{i=1}^{K_c}$

➜ The **clusters** are used to **estimate** the PWA separator

$$\min_{\{\omega_i,\gamma_i\}_{i=1}^{K_c}} \sum_{i=1}^{K_c} \sum_{C_i} \sum_{\substack{j=1\\ j\neq i}}^{K_c} \left\| \left(\begin{bmatrix} \chi(t)' & -1 \end{bmatrix} \begin{bmatrix} \omega_j - \omega_i \\ \gamma_j - \gamma_i \end{bmatrix} + 1 \right)_+ \right\|_2^2 + \lambda_p \sum_{i=1}^{K_c} (\|\omega_i\|_2^2 + \gamma_i^2),$$

⬇

$$\min_{\{\omega_i,\gamma_i\}_{i=1}^{K_c}} \sum_{i=1}^{K_c} \sum_{\substack{j=1\\ j\neq i}}^{K_c} \frac{1}{c_i} \left\| \left(\begin{bmatrix} X_i' & -\mathbf{1}_{c_i} \end{bmatrix} \begin{bmatrix} \omega_j - \omega_i \\ \gamma_j - \gamma_i \end{bmatrix} + \mathbf{1}_{c_i} \right)_+ \right\|_2^2 + \lambda_p \sum_{i=1}^{K_c} (\|\omega_i\|_2^2 + \gamma_i^2),$$

This **multicategory discrimination** problem can be **efficiently** solved through a regularized piecewise smooth Newton method with Armijo's line search (Breschi, Piga, Bemporad, 2016)

Step 2: illustrative example

$$x(t+1) = 0.8 \begin{bmatrix} \cos\alpha(t) & -\sin\alpha(t) \\ \sin\alpha(t) & \cos\alpha(t) \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t),$$

$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(t),$$

where

$$\alpha(t) = \begin{cases} \frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) \geq 0, \\ -\frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) < 0. \end{cases}$$

Chosen fixed-order controller structure

$$u(t) = u(t-1) + \theta_{s_c(t),1} e(t) + \theta_{s_c(t),2} e(t-1) + \theta_{s_c(t),3},$$

$$s_c(t) = i \iff \chi(t) = \begin{bmatrix} u(t-1) & y(t) & y(t-1) \end{bmatrix}' \in \mathcal{X}_i^c \text{ for } i = 1, 2$$

Step 2: illustrative example

$$x(t+1) = 0.8 \begin{bmatrix} \cos\alpha(t) & -\sin\alpha(t) \\ \sin\alpha(t) & \cos\alpha(t) \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t),$$
$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(t),$$

where

$$\alpha(t) = \begin{cases} \frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) \geq 0, \\ -\frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) < 0. \end{cases}$$

Chosen fixed-order controller structure

$$u(t) = u(t-1) + \theta_{s_c(t),1} e(t) + \theta_{s_c(t),2} e(t-1) + \theta_{s_c(t),3},$$
$$s_c(t) = i \iff \chi(t) = \begin{bmatrix} u(t-1) & y(t) & y(t-1) \end{bmatrix}' \in \mathcal{X}_i^c \text{ for } i = 1, 2$$

- **Estimate** the partition $\{\mathcal{X}_i^c\}_{i=1}^{K_c}$ *given* $\{\chi(t)\}_{t=1}^{T}$
- **Fix** the parameters of the local controllers and the mode sequence $\mathcal{S}_c = \{s_c(t)\}_{t=1}^{T}$

Step 3: update the mode sequence

For **fixed** parameters Θ and $\{\omega_i, \gamma_i\}_{i=1}^{K_c}$

$$\min_{\mathcal{S}_c} \mathcal{J}_1(\Theta^k, \mathcal{S}_c) + \mathcal{J}_2(\{\omega_i^k, \gamma_i^k\}_{i=1}^{K_c}, \mathcal{S}_c)$$

$$\mathcal{J}_1(\Theta^k, \mathcal{S}_c) = \sum_{t=1}^{T} \left(u(t) - X_c(\xi(t))' \theta^k_{s_c(t)} \right)^2$$

$$\mathcal{J}_2(\{\omega_i^k, \gamma_i^k\}_{i=1}^{K_c}, \mathcal{S}_c) = \sum_{t=1}^{T} \sum_{\substack{j=1 \\ j \neq s_c(t)}}^{K_c} \left\| \left(\begin{bmatrix} \chi(t)' & -1 \end{bmatrix} \begin{bmatrix} \omega_j^k - \omega^k_{s_c(t)} \\ \gamma_j^k - \gamma^k_{s_c(t)} \end{bmatrix} + 1 \right)_+ \right\|_2^2$$

- Use **dynamic programming**

Step 3: update the mode sequence

For **fixed** parameters Θ and $\{\omega_i, \gamma_i\}_{i=1}^{K_c}$

$$\min_{\mathcal{S}_c} \mathcal{J}_1(\Theta^k, \mathcal{S}_c) + \mathcal{J}_2(\{\omega_i^k, \gamma_i^k\}_{i=1}^{K_c}, \mathcal{S}_c)$$

$$\mathcal{J}_1(\Theta^k, \mathcal{S}_c) = \sum_{t=1}^{T} \left(u(t) - X_c(\xi(t))' \theta^k_{s_c(t)} \right)^2$$

$$\mathcal{J}_2(\{\omega_i^k, \gamma_i^k\}_{i=1}^{K_c}, \mathcal{S}_c) = \sum_{t=1}^{T} \sum_{\substack{j=1 \\ j \neq s_c(t)}}^{K_c} \left\| \left(\begin{bmatrix} \chi(t)' & -1 \end{bmatrix} \begin{bmatrix} \omega_j^k - \omega^k_{s_c(t)} \\ \gamma_j^k - \gamma^k_{s_c(t)} \end{bmatrix} + 1 \right)_+ \right\|_2^2$$

- Use **dynamic programming**

The mode sequence is updated by **trading off** between matching the measured input with a certain local controller and assigning χ to a certain polyhedral region

Step 3: illustrative example

$$x(t+1) = 0.8 \begin{bmatrix} \cos\alpha(t) & -\sin\alpha(t) \\ \sin\alpha(t) & \cos\alpha(t) \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t),$$
$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(t),$$

where

$$\alpha(t) = \begin{cases} \frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) \geq 0, \\ -\frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) < 0. \end{cases}$$

Chosen fixed-order controller structure

$$u(t) = u(t-1) + \theta_{s_c(t),1} e(t) + \theta_{s_c(t),2} e(t-1) + \theta_{s_c(t),3},$$
$$s_c(t) = i \iff \chi(t) = \begin{bmatrix} u(t-1) & y(t) & y(t-1) \end{bmatrix}' \in \mathcal{X}_i^c \text{ for } i = 1, 2$$

Step 3: illustrative example

$$x(t+1) = 0.8 \begin{bmatrix} \cos\alpha(t) & -\sin\alpha(t) \\ \sin\alpha(t) & \cos\alpha(t) \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t),$$
$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(t),$$

where

$$\alpha(t) = \begin{cases} \frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) \geq 0, \\ -\frac{\pi}{3} & \text{if } \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) < 0. \end{cases}$$

Chosen fixed-order controller structure

$$u(t) = u(t-1) + \theta_{s_c(t),1} e(t) + \theta_{s_c(t),2} e(t-1) + \theta_{s_c(t),3},$$
$$s_c(t) = i \iff \chi(t) = \begin{bmatrix} u(t-1) & y(t) & y(t-1) \end{bmatrix}' \in \mathcal{X}_i^c \text{ for } i = 1, 2$$

- **Estimate** the mode sequence $\mathcal{S}_c = \{s_c(t)\}_{t=1}^{T}$
- **Fix** the parameters of the local controllers and the partition $\{\mathcal{X}_i^c\}_{i=1}^{K_c}$

Input selection

Given:

- The estimated parameters $\Theta^\star$ and $\{\omega_i^\star, \gamma_i^\star\}_{i=1}^{K_c}$
- The current measured (noisy) closed-loop output $\tilde{y}(t)$
- The current reference signal $\tilde{r}(t)$
- The current vector $\tilde{\chi}(t)$ and regressor $\tilde{X}_c(t)$

Input selection

Given:

- The estimated parameters $\Theta^\star$ and $\{\omega_i^\star, \gamma_i^\star\}_{i=1}^{K_c}$
- The current measured (noisy) closed-loop output $\tilde{y}(t)$
- The current reference signal $\tilde{r}(t)$
- The current vector $\tilde{\chi}(t)$ and regressor $\tilde{X}_c(t)$

1. **find** the local controller to be used

$$\tilde{s}_c(t) \leftarrow \arg\max_{i=1,\ldots,K^c} (\omega_i^\star)' \tilde{\chi}(t) - \gamma_i^\star$$

2. **generate** the input $\tilde{u}(t)$ as

$$\tilde{u}(t) \leftarrow \tilde{X}_c(t)' \theta^\star_{\tilde{s}_c(t)}$$

Input selection

Given:

- The estimated parameters $\Theta^\star$ and $\{\omega_i^\star, \gamma_i^\star\}_{i=1}^{K_c}$
- The current measured (noisy) closed-loop output $\tilde{y}(t)$
- The current reference signal $\tilde{r}(t)$
- The current vector $\tilde{\chi}(t)$ and regressor $\tilde{X}_c(t)$

1. **find** the local controller to be used

$$\tilde{s}_c(t) \leftarrow \arg\max_{i=1,\ldots,K^c} (\omega_i^\star)'\tilde{\chi}(t) - \gamma_i^\star$$

2. **generate** the input $\tilde{u}(t)$ as

$$\tilde{u}(t) \leftarrow \tilde{X}_c(t)'\theta^\star_{\tilde{s}_c(t)}$$

The controller is **easy** to **deploy**

Outline

- Problem formulation
- Direct design of switching controllers from data
- A case study
- Concluding remarks

A case study

- In a real-world application,
  - ✓ $\mathcal{M}$ is usually given as a - rarely achievable - LTI model
  - ✓ The system dynamics is much more complex

A case study

- In a real-world application,
  - ✓ $\mathcal{M}$ is usually given as a - rarely achievable - LTI model
  - ✓ The system dynamics is much more complex
- A more realistic case study : Voltage-controlled DC motor with an unbalanced disk

A case study - cont'd

- Continuous-time system dynamics

$$\begin{bmatrix} \dot{\theta}(\tau) \\ \dot{\omega}(\tau) \\ \dot{I}(\tau) \end{bmatrix} = \left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & -\frac{b}{J} & \frac{K}{J} \\ 0 & -\frac{K}{L} & -\frac{R}{L} \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ \frac{mgl}{J} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{\sin(\theta(\tau))}{\theta(\tau)} \right) \begin{bmatrix} \theta(\tau) \\ \omega(\tau) \\ I(\tau) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix}$$

$$y(\tau) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta(\tau) \\ \omega(\tau) \\ I(\tau) \end{bmatrix},$$

A case study - cont'd

- Continuous-time system dynamics

$$\begin{bmatrix} \dot{\theta}(\tau) \\ \dot{\omega}(\tau) \\ \dot{I}(\tau) \end{bmatrix} = \left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & -\frac{b}{J} & \frac{K}{J} \\ 0 & -\frac{K}{L} & -\frac{R}{L} \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ \frac{mgl}{J} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{\sin(\theta(\tau))}{\theta(\tau)} \right) \begin{bmatrix} \theta(\tau) \\ \omega(\tau) \\ I(\tau) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix}$$

$$y(\tau) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta(\tau) \\ \omega(\tau) \\ I(\tau) \end{bmatrix},$$

- 3rd order system with dynamics depending on the load angular position θ

A case study - cont'd

- Continuous-time system dynamics

$$\begin{bmatrix} \dot{\theta}(\tau) \\ \dot{\omega}(\tau) \\ \dot{I}(\tau) \end{bmatrix} = \left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & -\frac{b}{J} & \frac{K}{J} \\ 0 & -\frac{K}{L} & -\frac{R}{L} \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ \frac{mgl}{J} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{\sin(\theta(\tau))}{\theta(\tau)} \right) \begin{bmatrix} \theta(\tau) \\ \omega(\tau) \\ I(\tau) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix}$$

$$y(\tau) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta(\tau) \\ \omega(\tau) \\ I(\tau) \end{bmatrix},$$

- 3rd order system with dynamics depending on the load angular position θ

- Model is **unknown** and we do not want to identify it

A case study - cont'd

- Continuous-time system dynamics

$$\begin{bmatrix} \dot{\theta}(\tau) \\ \dot{\omega}(\tau) \\ \dot{i}(\tau) \end{bmatrix} = \left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & -\frac{b}{J} & \frac{K}{J} \\ 0 & -\frac{K}{L} & -\frac{R}{L} \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ \frac{mgl}{J} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{\sin(\theta(\tau))}{\theta(\tau)} \right) \begin{bmatrix} \theta(\tau) \\ \omega(\tau) \\ I(\tau) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix}$$

$$y(\tau) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta(\tau) \\ \omega(\tau) \\ I(\tau) \end{bmatrix},$$

- 3rd order system with dynamics depending on the load angular position θ
- Model is **unknown** and we do not want to identify it
- We force $K_c = 2$ modes

A case study - cont'd

- Continuous-time system dynamics

$$\begin{bmatrix} \dot{\theta}(\tau) \\ \dot{\omega}(\tau) \\ \dot{I}(\tau) \end{bmatrix} = \left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & -\frac{b}{J} & \frac{K}{J} \\ 0 & -\frac{K}{L} & -\frac{R}{L} \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ \frac{mgl}{J} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{\sin(\theta(\tau))}{\theta(\tau)} \right) \begin{bmatrix} \theta(\tau) \\ \omega(\tau) \\ I(\tau) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix}$$

$$y(\tau) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta(\tau) \\ \omega(\tau) \\ I(\tau) \end{bmatrix},$$

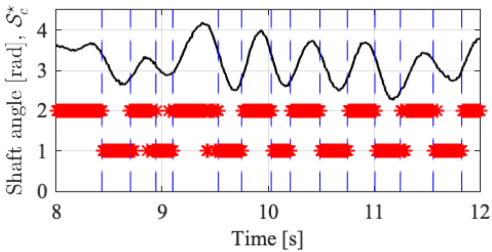
- Reference model

$$x_M(t+1) = 0.99 x_M(t) + 0.01 r(t)$$
$$y_d(t) = x_M(t),$$

A case study - cont'd

- 8th-order controller with $K_c = 2$ local controllers
- **Leverage** on LPV controller structure (Formentin, Piga, Tóth, Savaresi, 2016)
  - ➜ $\chi(t) = \begin{bmatrix}\theta(t-1) & \theta(t-2) & \theta(t-3) & \theta(t-4)\end{bmatrix}'$

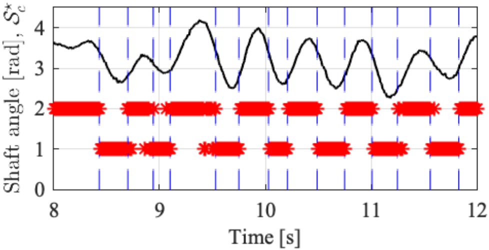
Training output and estimated mode sequence



A case study - cont'd

- 8th-order controller with $K_c = 2$ local controllers
- **Leverage** on LPV controller structure (Formentin, Piga, Tóth, Savaresi, 2016)

Training output and estimated mode sequence

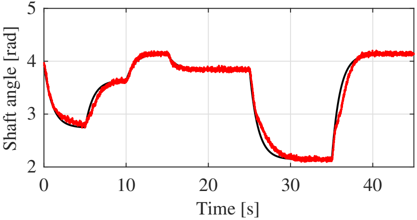


The controller switches when the shaft angle increases/decreases

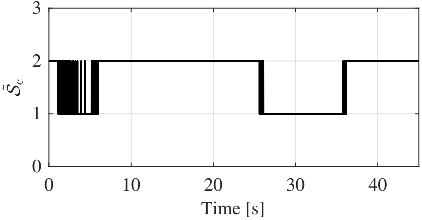
➜ The estimated controller is clearly linked to the physics of $\mathcal{G}_s$

A case study - cont'd

Desired vs attained closed-loop θ



Control mode sequence



- ✓ Competitive with LPV control
- ✓ More *robust* to noise (in the quasi-LPV case)
- ✓ PWA controller is trained 4x faster than LPV one

Outline

- Problem formulation
- Direct design of switching controllers from data
- A case study
- Concluding remarks

Concluding remarks

- New method for direct design of switching controllers from data
- Particularly suited for fixed-order controller design: no need for parameterization, identification, model/controller reduction
- May be preferred to data-driven LPV control: more robust and computationally more efficient
- Open problems: stability guarantees, performance bounds, constraints management

The End