Control-Oriented Learning

Simone Formentin*, Alessandro Chiuso+

*Department of Electronics, Information and Bioengineering
Politecnico di Milano (Italy)

+Department of Information Engineering
University of Padova (Italy)

Control Day
Padova, May 10, 2019

CoRe: Control-oriented Regularization

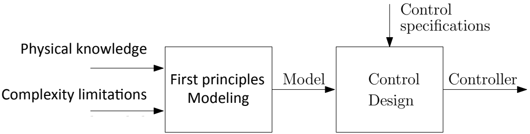
Simone Formentin*, **Alessandro Chiuso**+

*Department of Electronics, Information and Bioengineering
Politecnico di Milano (Italy)

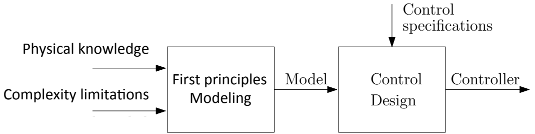
+Department of Information Engineering
University of Padova (Italy)

Control Day
Padova, May 10, 2019

Standard approach to control design

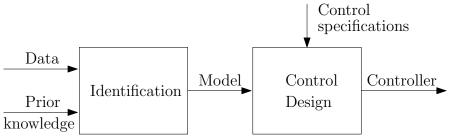


Standard approach to control design

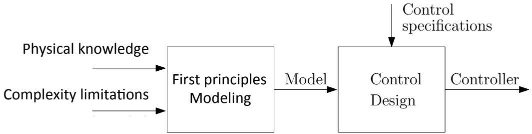


WARNING: Modeling is by far the most expensive step in a control project! ($\sim$ 75% of the total costs)

SYSID-based control design

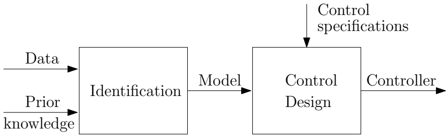


Standard approach to control design



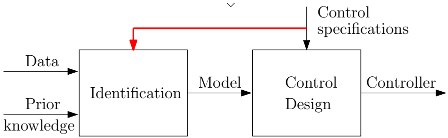
WARNING: Modeling is by far the most expensive step in a control project! ($\sim$ 75% of the total costs)

SYSID-based control design

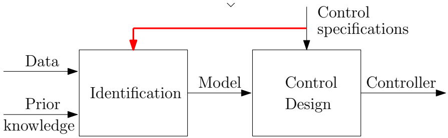


Two optimization steps: two objectives!

Identification for Control (I4C)



Identification for Control (I4C)



Model-Reference Control Objective

Choose C to **minimize**

$$V(C) = \left\| M - \frac{G_o C}{1 + G_o C} \right\|^2 ,$$

G_o = "true" unknown systems

Identification for Control (I4C)

- Control-oriented modeling step: find $\hat{G}$

$$J(G) = \left\| \frac{G_o C^*}{1 + G_o C^*} - \frac{G C^*}{1 + G C^*} \right\|^2$$

- (Nominal) control design step: find the new C^*

$$V(C) = \left\| M - \frac{\hat{G} C}{1 + \hat{G} C} \right\|^2$$

Identification for Control (I4C)

- Control-oriented modeling step: find $\hat{G}$

$$J(G) = \left\| \frac{G_o C^*}{1 + G_o C^*} - \frac{G C^*}{1 + G C^*} \right\|^2$$

- (Nominal) control design step: find the new C^*

$$V(C) = \left\| M - \frac{\hat{G} C}{1 + \hat{G} C} \right\|^2$$

Key trick: iterative procedure with closed-loop experiments using C^*

$$J_N(G) = \frac{1}{N} \sum_{t=1}^{N} (y^*(t) - G u^*(t))^2$$

Identification for Control (I4C)

- **Pro:** bias error distribution tuned for control design

Identification for Control (I4C)

- **Pro:** bias error distribution tuned for control design
- **Cons:**
  1. not clear how to account for priors;
  2. unbiased models might be a bad choice from a statistical perspective;
  3. no convergence to a local minimum of the control cost when the system is not in the model set.

Identification for Control (I4C)

- **Pro:** bias error distribution tuned for control design
- **Cons:**
  1. not clear how to account for priors;
  2. unbiased models might be a bad choice from a statistical perspective;
  3. no convergence to a local minimum of the control cost when the system is not in the model set.

- (Very) related area: **reinforcement learning**.

Identification for Control (I4C)

- **Pro:** bias error distribution tuned for control design
- **Cons:**
  1. not clear how to account for priors;
  2. unbiased models might be a bad choice from a statistical perspective;
  3. no convergence to a local minimum of the control cost when the system is not in the model set.

- (Very) related area: **reinforcement learning**.

- I4C can be reformulated as a **regularized identification problem**!

Control-oriented Regularization

Rationale

Decompose the model mismatch in $V(C) = \left\| M - \frac{G_o C}{1 + G_o C} \right\|^2$ as:

Control-oriented Regularization

Rationale

Decompose the model mismatch in $V(C) = \left\| M - \frac{G_o C}{1 + G_o C} \right\|^2$ as:

$$M - \frac{G_o C}{1 + G_o C} = \underbrace{M - \frac{GC}{1 + GC}}_{E_c} + \underbrace{\frac{GC}{1 + GC} - \frac{G_o C}{1 + G_o C}}_{E_m}$$

Decompose the model mismatch in $V(C) = \left\| M - \frac{G_o C}{1+G_o C} \right\|^2$ as:

$$M - \frac{G_o C}{1+G_o C} = \underbrace{M - \frac{GC}{1+GC}}_{E_c} + \underbrace{\frac{GC}{1+GC} - \frac{G_o C}{1+G_o C}}_{E_m}$$

- E_m is the modelling error (how close is G to G_o in closed loop)

Control-oriented Regularization

Rationale

Decompose the model mismatch in $V(C) = \left\| M - \frac{G_o C}{1 + G_o C} \right\|^2$ as:

$$M - \frac{G_o C}{1 + G_o C} = \underbrace{M - \frac{GC}{1 + GC}}_{E_c} + \underbrace{\frac{GC}{1 + GC} - \frac{G_o C}{1 + G_o C}}_{E_m}$$

- E_m is the modelling error (how close is G to G_o in closed loop)
- E_c is the control error using the model G

Control-oriented Regularization

Rationale

Decompose the model mismatch in $V(C) = \left\| M - \frac{G_o C}{1 + G_o C} \right\|^2$ as:

$$M - \frac{G_o C}{1 + G_o C} = \underbrace{M - \frac{GC}{1 + GC}}_{E_c} + \underbrace{\frac{GC}{1 + GC} - \frac{G_o C}{1 + G_o C}}_{E_m}$$

- E_m is the modelling error (how close is G to G_o in closed loop)
- E_c is the control error using the model G

Consider the upper bound

$$V(C) \leq U_V(G, C) = \alpha \left[\| E_m(G, C) \|^2 + \| E_c(G, C) \|^2 \right]$$

Control Oriented Regularization

Modeling Error

Recalling that **[Gianluigi's talk!]** identification in a regularization/Bayesian framework:

$$J(G) = \frac{1}{N}\sum_{t=1}^{N}(y_t - Gu_t)^2 + \|G\|_K^2$$

Recalling that **[Gianluigi's talk!]** identification in a regularization/Bayesian framework:

$$J(G) = \frac{1}{N}\sum_{t=1}^{N}(y_t - Gu_t)^2 + \|G\|_K^2$$

- $U_V(G,C)$ suggests to modify the fitting term

$$\frac{1}{N}\sum_{t=1}^{N}(y_t - Gu_t)^2$$

to keep $E_m = \frac{GC}{1+GC} - \frac{G_oC}{1+G_oC}$ small

Control Oriented Regularization

Modeling Error

Recalling that **[Gianluigi's talk!]** identification in a regularization/Bayesian framework:

$$J(G) = \frac{1}{N}\sum_{t=1}^{N}(y_t - Gu_t)^2 + \|G\|_K^2$$

- $U_V(G, C)$ suggests to modify the fitting term

$$\frac{1}{N}\sum_{t=1}^{N}(y_t - Gu_t)^2$$

to keep $E_m = \frac{GC}{1+GC} - \frac{G_oC}{1+G_oC}$ small

and the regularization term $\|G\|_K^2$ to keep $E_c = M - \frac{GC}{1+GC}$ small

$$E_m = \frac{GC}{1+GC} - \frac{G_oC}{1+G_oC} \simeq (G - G_o)\frac{C}{1+GC} = (G - G_o)W_C$$

where the last equation defines the weighting W_C

$$E_m = \frac{GC}{1+GC} - \frac{G_oC}{1+G_oC} \simeq (G-G_o)\frac{C}{1+GC} = (G-G_o)W_C$$

where the last equation defines the weighting W_C

Prediction error:

$$\lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N}(y_t - Gu_t)^2 = \|G-G_o\|^2_{\Phi_u} + \sigma^2$$

$$E_m = \frac{GC}{1+GC} - \frac{G_o C}{1+G_o C} \simeq (G - G_o)\frac{C}{1+GC} = (G - G_o)W_C$$

where the last equation defines the weighting W_C

Prediction error:

$$\lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N}(y_t - Gu_t)^2 = \|G - G_o\|^2_{\Phi_u} + \sigma^2$$

Therefore, with a suitable weighting filter ($W = W_C \Phi_u^{-1/2}$),

$$\frac{1}{N}\sum_{t=1}^{N}\|W(y_t - Gu_t)\|^2 \simeq \|G - G_o\|^2_{W_C W_C^*} + \sigma^2_{W_C}$$

$$E_c(G, C) = M - \frac{GC}{1 + GC}$$

is not linear (nor convex) in G, so we take its first-order Taylor expansion around the optimal solution

$$E_c^c(G, C) = M - (1 - M)GC$$

$$E_c(G, C) = M - \frac{GC}{1 + GC}$$

is not linear (nor convex) in G, so we take its first-order Taylor expansion around the optimal solution

$$E_c^c(G, C) = M - (1 - M)GC$$

The model should be regularized such that $E_c(G, C)$ is made small

Control-oriented Regularization

Combining the two costs

$$\|E_m(G,C)\|^2 + \|E_c(G,C)\|^2 \simeq \frac{1}{N}\sum_{t=1}^{N} \|W(y_t - Gu_t)\|^2 + \|E_c^c(G,C)\|^2$$

Control-oriented Regularization

Combining the two costs

$$\|E_m(G,C)\|^2 + \|E_c(G,C)\|^2 \simeq \frac{1}{N}\sum_{t=1}^{N}\|W(y_t - Gu_t)\|^2 + \|E_c^c(G,C)\|^2$$

Need to ensure stability

If we now include a penalty term (as usual) to guarantee stability of G, one should solve

$$\begin{aligned} \hat{G} &:= \min_G \|E_m(G,C)\|^2 + \|E_c(G,C)\|^2 + \|G\|_{\mathcal{H}}^2 \\ &\simeq \min_G \frac{1}{N}\sum_{t=1}^{N}\|W(y_t - Gu_t)\|^2 + \underbrace{\|E_c^c(G,C)\|^2}_{\textit{Quadratic in } G} + \|G\|_{\mathcal{H}}^2 \end{aligned}$$

"Ideally" would like to optimize

$$\hat{C} = \arg\min_{C \in \mathcal{C}} \underbrace{\left\| M - \frac{G_o C}{1 + G_o C} \right\|^2}_{V(C)},$$

"Ideally" would like to optimize

$$\hat{C} = \arg\min_{C \in \mathcal{C}} \underbrace{\left\| M - \frac{G_o C}{1 + G_o C} \right\|^2}_{V(C)},$$

Optimize least upper bound

$$\hat{C} = \arg\min_{C \in \mathcal{C}} \left[\min_{G} \left[\underbrace{\left\| \frac{GC}{1+GC} - \frac{G_o C}{1 + G_o C} \right\|^2}_{E_m} + \underbrace{\left\| M - \frac{GC}{1+GC} \right\|^2}_{E_c} \right] \right]$$

Control-oriented Regularization

Summary II

Optimize least upper bound

$$\hat{C} = \arg\min_{C \in \mathcal{C}} \left[\min_{G} \left[\underbrace{\left\| \frac{GC}{1+GC} - \frac{G_o C}{1+G_o C} \right\|^2}_{E_m} + \underbrace{\left\| M - \frac{GC}{1+GC} \right\|^2}_{E_c} \right] \right]$$

Control-oriented Regularization

Summary II

Optimize least upper bound

$$\hat{C} = \arg\min_{C \in \mathcal{C}} \left[\min_G \left[\underbrace{\left\| \frac{GC}{1+GC} - \frac{G_o C}{1+G_o C} \right\|^2}_{E_m} + \underbrace{\left\| M - \frac{GC}{1+GC} \right\|^2}_{E_c} \right] \right]$$

Optimize an approximation of the upper bound based on data

$$\hat{C} = \arg\min_{C \in \mathcal{C}} \left[\min_G \left[\underbrace{\sum_{t=1}^{N} \| W(y_t - G u_t) \|^2}_{\simeq E_m} + \underbrace{\| E_c^c(G) \|^2}_{\simeq E_c} \right] \right]$$

Control-oriented Regularization

Summary III

Optimize an approximation of the upper bound based on data

$$\hat{C} = \arg\min_{C \in \mathcal{C}} \left[\min_{G} \left[\underbrace{\sum_{t=1}^{N} \| W(y_t - G u_t) \|^2}_{\simeq E_m} + \underbrace{\| E_c^c(G) \|^2}_{\simeq E_c} \right] \right]$$

Control-oriented Regularization

Summary III

Optimize an approximation of the upper bound based on data

$$\hat{C} = \arg\min_{C \in \mathcal{C}} \left[\min_{G} \left[\underbrace{\sum_{t=1}^{N} \|W(y_t - Gu_t)\|^2}_{\simeq E_m} + \underbrace{\|E_c^c(G)\|^2}_{\simeq E_c} \right] \right]$$

Add regularization

$$\hat{C} = \arg\min_{C \in \mathcal{C}} \left[\min_{G} \left[\underbrace{\sum_{t=1}^{N} \|W(y_t - Gu_t)\|^2}_{\simeq E_m} + \underbrace{\|E_c^c(G)\|^2}_{\simeq E_c} + \underbrace{\|G\|_{\mathcal{H}}^2}_{\text{"ModelClass"}} \right] \right]$$

Control-oriented Regularization

Bayesian interpretation

$$\hat{C} := \arg\min_{C} \left[\min_{G} \frac{1}{N} \sum_{t=1}^{N} \| W(y_t - Gu_t) \|^2 + \underbrace{\lambda \| E_c^c(G, C) \|^2 + \| G \|_{\mathcal{H}}^2}_{\text{Quadratic in } G} \right]$$

Control-oriented Regularization

Bayesian interpretation

$$\hat{C} := \arg\min_{C} \left[\min_{G} \frac{1}{N} \sum_{t=1}^{N} \| W(y_t - G u_t) \|^2 + \underbrace{\lambda \| E_c^c(G, C) \|^2 + \| G \|_{\mathcal{H}}^2}_{\text{Quadratic in } G} \right]$$

with a **Gaussian Prior** for G

$$G \sim \mathcal{N}(m_C, K_C)$$

$$\begin{cases} K_C = (K^{-1} + \lambda T_w^\top T_w)^{-1} \\ m_C = \lambda K_C T_w^\top m \end{cases}$$

Control-oriented Regularization

Bayesian interpretation

$$\hat{C} := \arg\min_{C} \left[\min_{G} \frac{1}{N} \sum_{t=1}^{N} \| W(y_t - Gu_t) \|^2 + \underbrace{\lambda \| E_c^c(G, C) \|^2 + \| G \|_{\mathcal{H}}^2}_{\textit{Quadratic in G}} \right]$$

with a **Gaussian Prior** for G

$$G \sim \mathcal{N}(m_C, K_C)$$

$$\begin{cases} K_C = (K^{-1} + \lambda T_w^\top T_w)^{-1} \\ m_C = \lambda K_C T_w^\top m \end{cases}$$

(K_C is the **control-oriented kernel**), translates into

$$\hat{C} := \arg\min_{C} \left[\min_{G} \frac{1}{N} \sum_{t=1}^{N} \| W(y_t - Gu_t) \|^2 + \| G - m_C \|_{K_C^{-1}}^2 \right]$$

Control-oriented Regularization

Bayesian interpretation

$$\hat{C} := \arg\min_{C} \left[\min_{G} \frac{1}{N} \sum_{t=1}^{N} \| W(y_t - G u_t) \|^2 + \underbrace{\lambda \| E_c^c(G, C) \|^2 + \| G \|_{\mathcal{H}}^2}_{\text{Quadratic in } G} \right]$$

with a **Gaussian Prior** for G

$$G \sim \mathcal{N}(m_C, K_C)$$

$$\begin{cases} K_C = (K^{-1} + \lambda T_w^\top T_w)^{-1} \\ m_C = \lambda K_C T_w^\top m \end{cases}$$

(K_C is the **control-oriented kernel**), translates into

$$\hat{C} := \arg\min_{C} \left[\min_{G} \frac{1}{N} \sum_{t=1}^{N} \| W(y_t - G u_t) \|^2 + \| G - m_C \|_{K_C^{-1}}^2 \right]$$

Hyperparameters found via **marginal likelihood** maximization

$$p_{\lambda,C}(Y) = \int_{\mathcal{H}} p_{\lambda,C}(Y, G)\, dG \quad \propto \int_{\mathcal{H}} p_{\lambda,C}(Y|G) p_{\lambda,C}(G)\, dG$$

Control-oriented Regularization

Marginal Likelihood

$$p_{\lambda,C}(Y) = \int_{\mathcal{H}} p_{\lambda,C}(Y,G)\,dG \quad \propto \int_{\mathcal{H}} p_{\lambda,C}(Y|G)p_{\lambda,C}(G)\,dG$$

Log - Likelihood

$$\log p_{\lambda,C}(Y) \quad \propto \quad \log \int_{\mathcal{H}} e^{-\frac{1}{N}\sum_{t=1}^{N} \|W(y_t - Gu_t)\|^2 + \|G - m_C\|^2_{K_C^{-1}}} dG + \log\det K_C + const$$

Control-oriented Regularization

Marginal Likelihood

$$p_{\lambda,C}(Y) = \int_{\mathcal{H}} p_{\lambda,C}(Y,G)\, dG \quad \propto \int_{\mathcal{H}} p_{\lambda,C}(Y|G) p_{\lambda,C}(G)\, dG$$

Log - Likelihood

$$\log p_{\lambda,C}(Y) \quad \propto \quad \log \int_{\mathcal{H}} e^{-\frac{1}{N}\sum_{t=1}^{N} \|W(y_t - Gu_t)\|^2 + \|G - m_C\|^2_{K_C^{-1}}} dG \\ + \log\det K_C + const$$

Can be used to

$$p_{\lambda,C}(Y) = \int_{\mathcal{H}} p_{\lambda,C}(Y,G)\,dG \quad \propto \int_{\mathcal{H}} p_{\lambda,C}(Y|G)p_{\lambda,C}(G)\,dG$$

Log - Likelihood

$$\log p_{\lambda,C}(Y) \quad \propto \quad \log \int_{\mathcal{H}} e^{-\frac{1}{N}\sum_{t=1}^{N}\|W(y_t - Gu_t)\|^2 + \|G - m_C\|^2_{K_C^{-1}}} dG + \log\det K_C + const$$

Can be used to

- Optimize hyperparameters (e.g. λ)

$$p_{\lambda,C}(Y) = \int_{\mathcal{H}} p_{\lambda,C}(Y, G)\, dG \quad \propto \int_{\mathcal{H}} p_{\lambda,C}(Y|G) p_{\lambda,C}(G)\, dG$$

Log - Likelihood

$$\log p_{\lambda,C}(Y) \quad \propto \quad \log \int_{\mathcal{H}} e^{-\frac{1}{N}\sum_{t=1}^{N} \|W(y_t - Gu_t)\|^2 + \|G - m_C\|^2_{K_C^{-1}}} dG + \log\det K_C + const$$

Can be used to

- Optimize hyperparameters (e.g. λ)
- Understanding whether a certain C (e.g. at the $k-th$ iteration of the algorithm, see next slide) is good enough (stopping criteria)

Control-oriented Regularization

The algorithm

CoRe identification procedure

Iterate (until convergence):

1. $$\hat{G}^{(k)} = \arg\min_{G} \left[\frac{1}{N} \sum_{t=1}^{N} \| W^{(k-1)}(y_t - G u_t) \|^2 + \| G - m_{\hat{C}^{(k-1)}} \|^2_{K^{-1}_{\hat{C}^{(k-1)}}} \right]$$

Control-oriented Regularization

The algorithm

CoRe identification procedure

Iterate (until convergence):

1. $$\hat{G}^{(k)} = \arg\min_{G} \left[\frac{1}{N} \sum_{t=1}^{N} \| W^{(k-1)}(y_t - G u_t) \|^2 + \| G - m_{\hat{C}^{(k-1)}} \|^2_{K^{-1}_{\hat{C}^{(k-1)}}} \right]$$
2. $$\hat{C}^{(k)} = \arg \min_{C} \left\| M - \frac{\hat{G}^{(k)} C}{1 + \hat{G}^{(k)} C} \right\|^2$$

CoRe identification procedure

Iterate (until convergence):

1. $$\hat{G}^{(k)} = \arg\min_{G} \left[\frac{1}{N} \sum_{t=1}^{N} \|W^{(k-1)}(y_t - G u_t)\|^2 + \|G - m_{\hat{C}^{(k-1)}}\|^2_{K^{-1}_{\hat{C}^{(k-1)}}} \right]$$
2. $$\hat{C}^{(k)} = \arg\min_{C} \left\| M - \frac{\hat{G}^{(k)} C}{1 + \hat{G}^{(k)} C} \right\|^2$$

Proposition: The algorithm converges to a local optimum of the upper bound of $V(C)$.

Control-oriented Regularization

Alternative to step 2 [with G. Pillonetto and A. Scampicchio]

"Cautious" CoRe identification procedure

Iterate (until convergence):

1. $$\hat{G}^{(k)} = \arg\min_{G} \left[\frac{1}{N} \sum_{t=1}^{N} \| W^{(k-1)}(y_t - G u_t) \|^2 + \| G - m_{\hat{C}^{(k-1)}} \|^2_{K^{-1}_{\hat{C}^{(k-1)}}} \right]$$

Control-oriented Regularization

Alternative to step 2 [with G. Pillonetto and A. Scampicchio]

"Cautious" CoRe identification procedure

Iterate (until convergence):

1. $$\hat{G}^{(k)} = \arg\min_{G} \left[\frac{1}{N} \sum_{t=1}^{N} \| W^{(k-1)}(y_t - Gu_t)\|^2 + \|G - m_{\hat{C}^{(k-1)}}\|^2_{K^{-1}_{\hat{C}^{(k-1)}}} \right]$$
2. $$\hat{C}^{(k)} = \arg\min_{C} \mathbb{E}_{p^{(k)}(G|Y)} \left[\underbrace{\|E_c^c(G)\|^2}_{Quadratic \ \ in \ \ G} |Y \right]$$

Example

Consider a linear model

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

Example

Consider a linear model

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

the class of controllers:

$$C(z,\rho) = \frac{\rho_0 + \rho_1 z^{-1} + \rho_2 z^{-2} + \rho_3 z^{-3} + \rho_4 z^{-4} + \rho_5 z^{-5}}{1 - z^{-1}},$$

Example

Consider a linear model

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

the class of controllers:

$$C(z,\rho) = \frac{\rho_0 + \rho_1 z^{-1} + \rho_2 z^{-2} + \rho_3 z^{-3} + \rho_4 z^{-4} + \rho_5 z^{-5}}{1 - z^{-1}},$$

the achievable and unachievable objectives

$$M_{\mathrm{a}} = \frac{C_o G_o}{1 + C_o G_o}, \qquad M_{\mathrm{u}}(z) = \frac{(1-\theta)^2 z^{-3}}{1 - 2\theta z^{-1} + \theta^2 z^{-2}}, \quad \theta = e^{-10 T_s}$$

Example

Identified system (achievable case)

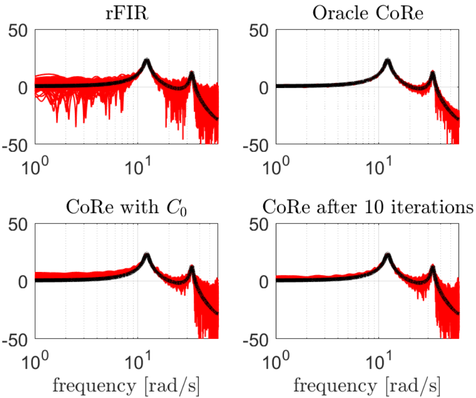


Figure: Magnitude of the frequency responses of the plant G_o (black line) and the identified model $\hat{G}$ for different realizations of the output noise (red lines).

Example

Closed-loop system (achievable case)

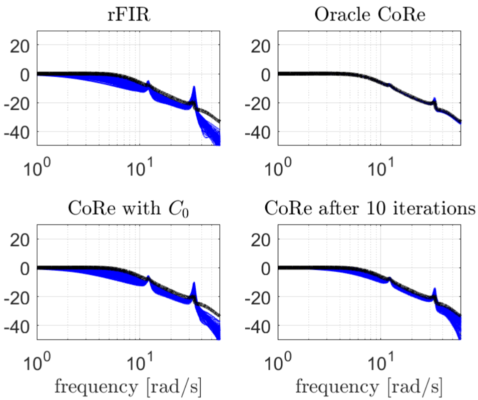


Figure: Magnitude of the frequency responses of the reference model (black line) and the closed-loop system F obtained using the model identified with different realizations of the output noise (blue lines).

Example

Closed-loop performance (achievable case)

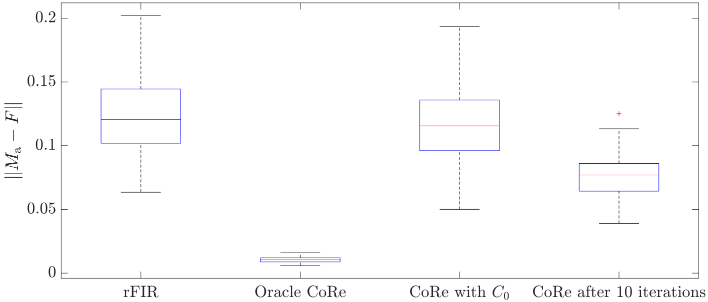


Figure: Boxplot of the closed-loop matching errors between M_a and F in the 4 considered scenarios.

Example

Closed-loop performance (unachievable case)

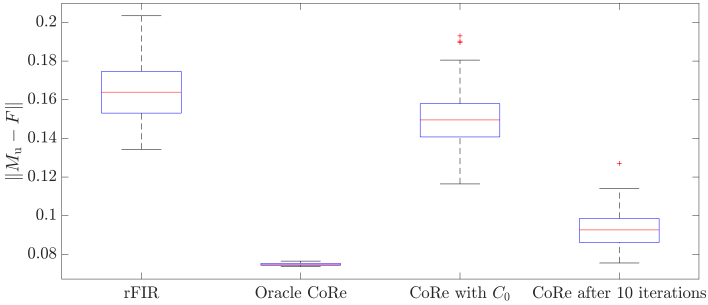


Figure: Boxplot of the closed-loop matching errors between M_u and F in the 4 considered scenarios.

Example

Classical PEM-based I4C

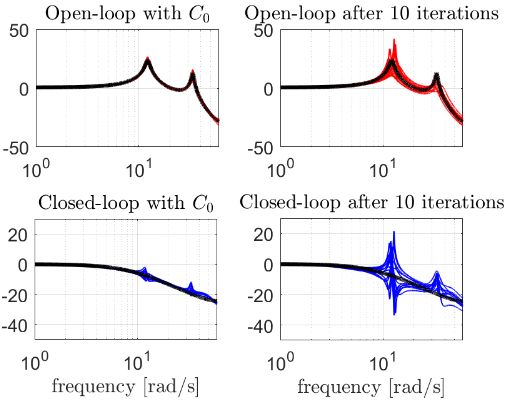


Figure: Magnitude of the frequency responses of the ideal open and closed-loop models (black line) as compared to the closed-loop system (blue line) obtained using the identified model (red line). Model order assumed to be known!

Example

Number of iterations

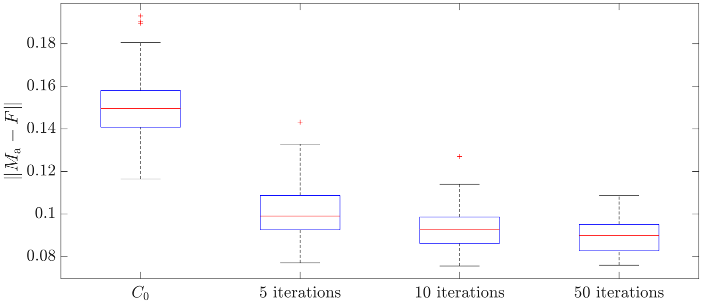


Figure: Sensitivity to the number of iterations of the closed-loop matching errors between M_a and F using the CoRe approach.

Example

A tentative stopping criterion: look at the marginal likelihood!

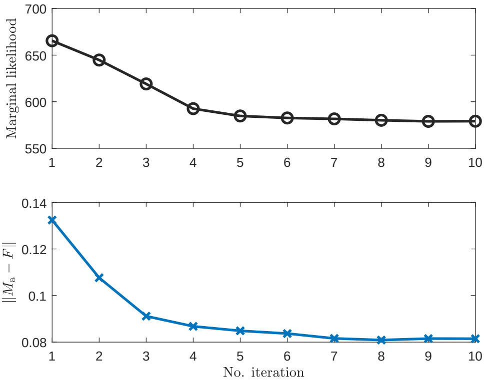


Figure: Marginal likelihood along iterations as compared to the closed-loop model matching cost.

Stability (with high probability)

Stability (with high probability)

Stability

Let $\varepsilon_s \in (0,1)$, $\delta_s \in (0,1)$ be assigned probability levels. Draw $N_s = \lceil \ln(2/\delta_s)/2\varepsilon_s^2 \rceil$ models $\hat{G}_k$, $k = 1, \ldots, N_s$, from the $p(G|Y)$. If the controller stabilizes all the $\hat{G}_k$'s, then

$$\mathbb{P}\left(p \geq 1 - \varepsilon_s\right) \geq 1 - \delta_s. \qquad p = \mathbb{P}[C \;\; stabilizes \;\; G]$$

Performance (with high probability)

Performance (with high probability)

Performance

Let $\varepsilon_p \in (0,1)$, $\delta_p \in (0,1)$ be assigned probability levels. Draw $N_p = \lceil \ln(1/\delta_p)/\ln(1/(1-\varepsilon_p)) \rceil$ models $\hat{G}_k$, $k = 1, \ldots, N_p$, from $p(G|Y)$ and compute the sample worst case performance

$$\hat{V}_{wc} = \max_{k=1,\ldots,N_p} \left\| M - \frac{\hat{G}_k C}{1 + \hat{G}_k C} \right\|^2 \tag{1}$$

Then, with confidence $1 - \delta_s$, it holds that

$$\mathbb{P}\left(V_{wc} \geq \hat{V}_{wc} \right) \leq \varepsilon_p.$$

Conclusions

- Control-oriented modeling from data should be seen as a **regularized identification problem**.

Conclusions

- Control-oriented modeling from data should be seen as a **regularized identification problem**.
- Regularization should not only take into account priors on system dynamics, but also **priors on model application** (i.e., desired control performance).

Conclusions

- Control-oriented modeling from data should be seen as a **regularized identification problem**.
- Regularization should not only take into account priors on system dynamics, but also **priors on model application** (i.e., desired control performance).
- CoRe turns out to be an iterative procedure based on **one set of I/O data only**.

Conclusions

- Control-oriented modeling from data should be seen as a **regularized identification problem**.
- Regularization should not only take into account priors on system dynamics, but also **priors on model application** (i.e., desired control performance).
- CoRe turns out to be an iterative procedure based on **one set of I/O data only**.
- Marginal likelihood provides an indication on the termination criterion.

Conclusions

- Control-oriented modeling from data should be seen as a **regularized identification problem**.
- Regularization should not only take into account priors on system dynamics, but also **priors on model application** (i.e., desired control performance).
- CoRe turns out to be an iterative procedure based on **one set of I/O data only**.
- Marginal likelihood provides an indication on the termination criterion.
- In the tested example, the obtained performance with a few iterations is close to the oracle.