Limited-Complexity Controller Tuning: A Set Membership Data-Driven Approach

Freddy Valderrama[a,b], Fredy Ruiz[a,c,*]

[a] *Departamento de Electrónica, Pontificia Universidad Javeriana, Carrera 7 No. 40 - 62, Bogotá, Colombia*
[b] *Escuela De Ciencias Básicas Tecnología e Ingeniería, UNAD, Bogotá, Colombia*
[c] *Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Via Ponzio 34/5, 20133 Milano, Italy*

Abstract

Data-driven tuning is an alternative to model-based controller design where controllers are directly identified from data, avoiding a plant identification step. In this paper, an approach to tune limited-complexity controllers from data for linear systems is proposed. The controller is parametrized as a linear combination of a large set of basis functions and the proposed algorithm allows to select a sparse subset of bases, guaranteeing a bounded approximation error. A feasibility condition allows to adjust the trade-off between accuracy and sparsity. The controller design is performed by solving a set of linear programming problems, allowing to handle large data-sets. The proposed strategy is evaluated by means of a Monte-Carlo simulation experiment on a flexible transmission benchmark model. Results show that the proposed solution offers similar results than previous approaches for large data-sets, requiring less adjustable parameters. However, for reduced data-sets, the presented algorithm shows better performance than the compared approaches.

Keywords: Identification for control, Uncertain systems, Linear systems, Model/Controller reduction.

1. Introduction

In a standard approach to model-based controller design (MB), two optimization problems are to be solved. First, a system model is built (either from first principles, estimated from experimental data or a mix of both). Secondly, a controller is selected or tuned, which is possibly optimal in some sense, with respect to the built model. Sometimes an order reduction step of the resulting controller is required.

*Corresponding author

Email addresses: `f.valderrama@javeriana.edu.co` (Freddy Valderrama), `fredy.ruiz@polimi.it` (Fredy Ruiz)

The performance of the derived controller relies, on great measure, on the quality of the employed model. Note that, obtaining reliable models of industrial processes turns out to be a complex task and their construction might involve excessive costs [1]. Data-driven control (DDC) approaches do not rely on an estimated plant model, since the available input-output data, experimentally collected from the plant, is directly used to design the controller, reducing the design problem to a single optimization step. A theoretical comparison between MB and DDC design is reported in [2].

In a linear framework, sequential and batch methods have been developed to deal with the DDC problem. In sequential methods, a series of experiments are required to adjust the controller parameters, whereas, in batch methods a single experiment allows to obtain the controller parameters. Some of the most known batch methods are Virtual Reference Feedback Tuning (VRFT) and Non-iterative Correlation based Tuning (CbT).

The VRFT method is developed in an stochastic framework and uses instrumental variables as the identification method employed to solve the controller identification problem. In the presence of noisy measurement, a second experiment or a plant model is required, see e.g., [3]. The VRFT method has been extended to non-linear systems in [4]. An extension to MIMO linear systems has been developed in [5]. Some applications can be found in [6, 7, 8].

The CbT method in [9] follows also a stochastic approach, where a properly constructed signal is required to deal with noisy measurements. In a real setting such signal might be difficult to obtain. Satisfactory applications and extensions of these methods have been reported. In [10], closed loop stability conditions for the CbT method have been proposed.

Set Membership (SM) estimation techniques, where noises are assumed as unknown-but-bounded signals, have been satisfactorily applied in system identification [11, 12], filter design from data [13, 14, 15] and controller design from data [16, 17, 18, 19]. The problem of data-driven controller tuning for linear systems has been investigated in [16, 17]. In [17], the SM Errors-in-Variables (SMEiV) identification method is applied to solve the controller tuning problem. Convex relaxations are employed to solve the resulting polynomial optimization problems, leading to computationally demanding solutions and therefore limiting the amount of experimental data that can be considered, even for a reduced set of controller parameters. In [16], set over-bounding techniques are used to derive efficient linear programming problems from the original non-convex problem, allowing to manage larger data sets. In the case of non linear systems, the controller design from data problem has been tackled in [18, 19], where all state variables are assumed to be measured and a feasible state trajectory is required as reference signal, generated for example by an expert human operator.

As stated in [20], in industrial practice low complexity controllers are preferred. The main problems related to high complexity controllers implementation are essentially two:

- When the number of controller parameters is large, many arithmetic operations, i.e., multiplications and additions,are required, thus slowing down the computational processing.

- High-complexity controllers are fragile, i.e. highly sensitive to round-off errors.

In literature, low complexity controllers design has been addressed by means of different approaches. In [21], techniques to derive low order controllers by previously performing model order-reduction are exposed. Another alternative is shown in [22] where the high-complexity controllers derived from high-complexity models can be approximated with low-complexity ones by means of controller order-reduction. In [23], fixed-order controllers are tuned from high-complexity models without explicit order-reduction. All the previous approaches are model-based, requiring a mathematical model of the process.

A data-driven approach to reduced-order controller design was proposed in [20]. Specifically, an iterative algorithm based on CbT and ℓ_1 regularization to design sparse-controllers was proposed. The framework assumes stochastic noises and the solution is based on an iterative process that can present convergence problems.

In this paper, an alternative DDC approach to design low-complexity controllers using a Set Membership estimation framework is proposed. The main contributions are:

- An efficient algorithm to derive sparse controllers from data is proposed. Linear and quadratic programming programs are employed to estimate optimal controller parameters, avoiding polynomial problems.
- The problem of noisy measurements is addressed without statistical assumptions on the disturbance signals, overcoming the limitations suffered by existing statistical solutions when reduced data sets are available.
- The proposed controller tuning algorithm does not require iterations or multiple experiments. Moreover, a criterion is provided to suitably manage the trade-off between accuracy and sparsity.

The outline of the paper is as follows. In Section II, the problem formulation is presented. In Section III, a Set Membership framework to sparse-controller tuning is comprehensively described. Finally, in Section IV the proposed solution is evaluated on a benchmark problem, comparing its performance with an existing approach using sparse CbT. The conclusions end the paper in Section V.

2. Problem formulation

In this Section the main problem is formulated. First, the setting and main assumptions are presented.

Consider a discrete-time linear-time invariant (LTI) single-input single-output (SISO) feedback control scheme, as depicted in Fig. 1, where q^{-1} denotes the

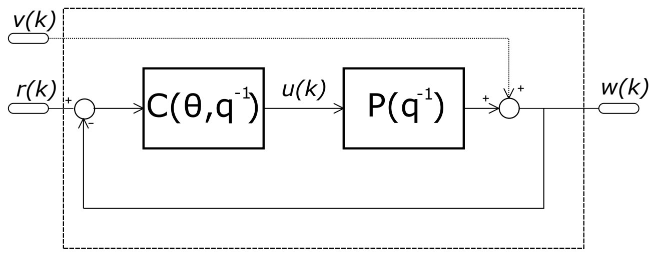

Figure 1: Assumed feedback control structure

backward shift operator, $P(q^{-1})$ is the plant transfer function, $C(\theta, q^{-1})$ is the controller transfer function, θ is a vector of controller parameters, $r(k)$ is the reference signal, $v(k)$ is output noise, $u(k)$ and $w(k)$ are the plant input and output signals, respectively. For the system interconnection in Fig. 1, the aim of the controller tuning procedure is to select an optimal controller C^o minimizing some performance criterion. For example, an optimization problem can be stated as:

$$C^o(\theta^o, q^{-1}) = \arg \min_{C(\theta,q^{-1}) \in \mathcal{C}} J(C(\theta, q^{-1})) \tag{1}$$

For the cost function

$$J(\theta) = \left\| M(q^{-1}) - \frac{P(q^{-1})C(\theta, q^{-1})}{1 + P(q^{-1})C(\theta, q^{-1})} \right\|_{\mathfrak{s}} \tag{2}$$

Being $\mathfrak{s}$ a proper system norm, $\mathcal{C}$ the set of LTI systems where the controller is selected, θ a real vector parameterizing $C(\theta, q^{-1})$ and $M(q^{-1})$ a reference model for the closed-loop system, where performance specifications are embedded. If $P(q^{-1})$ is known, Problem 1 can be seen as a loop-shaping problem and, for $\mathcal{H}_2$, $\mathcal{H}_\infty$ and ℓ_1 norms, known techniques exist to solve it under proper controllability (reachability) and observability (detectability) conditions. See e.g., [24],[25].

If system $P(q^{-1})$ is unknown, Problem (1) can not be solved directly. The common procedure to controller design for unknown plants is to follow a two-step procedure where first a system model $\hat{P}(q^{-1})$ is estimated from data, possibly with some uncertainty model and then, a controller is obtained solving problem (1) for $\hat{P}(q^{-1})$.

In the framework proposed in this paper, the controller is parametrized as a linear combination of fixed basis functions, that is,

$$C(\theta, q^{-1}) = \sum_{i=1}^{m_{max}} \theta_i \beta_i(q^{-1}) \tag{3}$$

Where m_{max} is designated as the maximum allowable controller complexity.

Then, the controller that solves problem (1) is selected from the set:

$$\mathcal{C} = \left\{ C(\theta, q^{-1}) : \theta \in \Theta \subseteq \Re^{m_{max}} \right\}$$

The following assumptions define the framework of the data-driven controller tuning problem.

Assumption 1. *$P(q^{-1})$ is unknown. The available information on $P(q^{-1})$ is a set of noisy input-output data generated by $P(q^{-1})$, initially at rest,*

$$\mathcal{D} = \{w(k), u(k), k = 1, 2, ..., N\} \tag{4}$$

Where

$$w(k) = \sum_{j=0}^{k} h_j^P u(k-j) + v(k),$$

h_j^P are the impulse response coefficients of $P(q^{-1})$ and $v(k)$ is the plant output noise/disturbance.

Remark 1. *Input $u(t)$ is assumed sufficiently informative, i.e., it allows to obtain bounded sets of controller parameters. However, no hypotheses are established about the plant operation during the experiment. Either open or closed loop data can be employed.*

Assumption 2. *An internally stabilizing controller $C(\theta^0, q^{-1})$ exists such that the minimum of* (2) *is* 0*, that is,*

$$M(q^{-1}) = \frac{P(q^{-1})C(\theta^0, q^{-1})}{1 + P(q^{-1})C(\theta^0, q^{-1})}$$

for some $\theta^o \in \Theta$.

The previous assumption is required for the derivation of theoretical results of the controller tuning method proposed here.

Assumption 3. *The optimal controller $C(\theta^0, q^{-1})$ is sparse, in other words*

$$\left\| \theta^0 \right\|_0 \ll m_{max}$$

where

$$\|\theta\|_0 = card(supp(\theta))$$

$$supp(\theta) \doteq \{i \in \{1, 2, ..., m_{max}\}, \theta_i \neq 0\}$$

and $card(.)$ is the set cardinality.

Assumption 3 is reasonable, because one would expect that from a "large" set of basis functions, only some of them are useful for solving problem (1), if the set $\mathcal{C}$ is properly parametrized.

Assumption 4. *Any controller $C(\theta, q^{-1}) \in \mathcal{C}$ can be expressed as:*

$$C(\theta, q^{-1}) = C^o(\theta^o, q^{-1})\left(1 + \Delta(\theta, q^{-1})\right)$$

with $\Delta(\theta, q^{-1})$ a proper and stable transfer function.

The previous assumption is motivated by the fact that common controller structures are usually stable or marginally stable (with pure integrators). Then, if C^o and C share poles in $z = 1$ with the same multiplicity, the difference between them is always a stable system.

The noise sequence $v(k)$ is modeled as an unknown but bounded (UBB) signal, without any statistical assumption about it, as follows:

Assumption 5. *Noise $v(k)$ is an (UBB) signal, such that*

$$\|v(k))\|_{\ell_p} \leq \epsilon_p$$

with $\ell_p \in \{2, \infty\}$. In this framework, energy and amplitude limited noise sequences can be considered.

Based on the previous assumption, the data-driven controller tuning problem can be stated as follows:

Problem 1. Sparse-controller tuning: *Given a data set $\mathcal{D}$, generated as in Assumptions 1 and 5; a reference model $M(q^{-1})$ and a set of basis functions $\beta_i, i = 1, ..., m_{max}$ satisfying assumptions 2, 3 and 4 identify a coefficient vector θ such that*

1. *θ is sparse*
2. *$J(\theta)$ in (2) is "small"*

Note that, since $J(\theta)$ depends on the plant but according to Assumption 1 it is unavailable, it is necessary to express the cost function in terms of the available data.

3. A sparse Set Membership framework for controller tuning

Departing from the assumptions stated in the previous section, in this section, Problem 1 is cast into a Set Membership identification framework and an efficient algorithm is proposed to find sparse controllers from data.

Let the model matching error be defined as the argument of the cost function $J(\theta)$ in (2):

$$E_m(\theta) = M - \frac{PC(\theta)}{1 + PC(\theta),} \tag{5}$$

where the shift operator q^{-1} has been removed for simplicity. This notation is maintained in the following when possible.

First, the model matching error is expressed in a convenient form.

Theorem 1. *For any controller $C(\theta, q^{-1}) \in \mathcal{C}$, satisfying assumption 4, the model matching error $E_m(\theta)$ in (2) can be expressed as:*

$$E_m(\theta) = \frac{1}{1 + M\Delta(\theta)} \left(M(1 - M) - C(\theta)(1 - M)^2 P \right) \tag{6}$$

Proof: See Appendix 5.

Fig. 2 shows a block-diagram of the equivalent model matching error system, derived in Theorem 1. Note that the system contains an output inverse multiplicative uncertainty structure that allows to state the following Corollary:

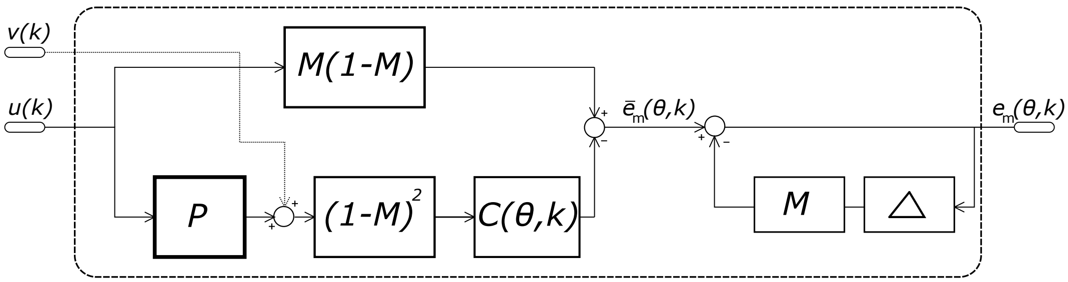

Figure 2: Model Matching block diagram

Corollary 1. *Given a controller $C(\theta) \in \mathcal{C}$, the transfer function $E_m(\theta)$ is input-output stable if:*

$$|\Delta(\theta, e^{j\omega})| < |M(e^{j\omega})|, \forall\omega \in [0,\ \pi] \tag{7}$$

The result follows from the Nyquist stability condition for the inverse multiplicative uncertainty structure in eq. (6).

The following Corollary allows to transform the model-based controller design problem in eq. (1) into an identification problem.

Corollary 2. *Given a data set $\mathcal{D}$ generated as in assumption 1, affected by noise bounded as in assumption 5, any controller $C(\theta) \in \mathcal{C}$, guaranteeing an internally stable loop, satisfies the time-domain relations:*

$$e_m(\theta, k) = \frac{1}{1 + M\Delta(\theta)} * \overline{e}_m(k) \tag{8}$$

$$\begin{aligned} \overline{e}_m(\theta, k) &= [M(1 - M)] * u(k) - \\ &\quad [C(\theta)\,(1 - M)^2] * w(k) + d(\theta, k) \end{aligned} \tag{9}$$

$$d(\theta, k) = (1 + \Delta(\theta))C^o\,(1 - M)^2 * v(k) \tag{10}$$

and $d(\theta, t)$ is an UBB signal with bound

$$\|d(\theta, k)\|_{\ell_p} \le \left\| (1 + \Delta(\theta))C^o\,(1 - M)^2 \right\|_{\ell_p, \ell_p} \epsilon_p = \delta_p \tag{11}$$

Moreover, for an optimal controller $C^o(\theta^o)$, it holds:

$$\begin{array}{rcl} e_m(k) & = & \overline{e}_m(k) \\ M(1-M)*u(k)+d(\theta^o,k) & = & C^o(\theta^o)(1-M)^2*w(k) \end{array} \tag{12}$$

Proof: The results are the time domain relations of applying the signal $u(k)$ to the model matching error system in (2) and using the experimentally measured plant output $w(k)$ as input to the sub-system $(1-M)^2$ in the lower branch of the block-diagram in Fig. 2.

The resulting output noise $d(\theta,k)$ is bounded because the transfer function in (10) is input-output stable under the hypotheses that $\Delta(\theta)$ is stable and $C^o(\theta^o)$ guarantees an internally stable loop.

From the previous development, it is possible to cast the data-driven controller tuning problem into an identification problem.

let

$$y_c(k) = M(1-M)*u(k) \tag{13}$$

$$u_c(k) = (1-M)^2\, w(k) \tag{14}$$

Then, for an optimal controller $C(\theta^0)$, eq. (12) can be rewritten as

$$y_c(k)+d(\theta^o,k)=C(\theta^0)u_c(k) \tag{15}$$

Note that, equation (15) corresponds to an identification problem with additive noise, where the system to be estimated is $C(\theta^0)$.

While the available data set is generated within an errors-in-variables setting, with the unknown signal $v(k)$ affecting the controller input, in this work, it is proposed to perform an over-bounding of the noise, moving it to the controller output, thanks to the guaranteed output noise bound established in (11). Previous approaches have used errors-in-variables identification algorithms, that in a Set-Membership setting lead to highly complex optimization problems, see e.g. [17].

3.1. The Feasible Parameters Set

From the previous analysis, we are in position to define the Feasible Parameters Set (FPS), that is, the set of all controller parameters θ that are compatible with hypotheses and data.

Definition 1. *Feasible Parameter Set (FPS):*

$$FPS=\left\{\theta\in\Theta:\|Y_c-\Phi\theta\|_{\ell_p}\leq\delta_p\right\} \tag{16}$$

where

$$Y_c=[y_c(1),\ y_c(2),\dots y_c(N)]^T,$$

$$\Phi = [\phi^1,\ \phi^2, \dots \phi^{m_{max}}]$$

and

$$\phi^i = [y_{\beta i}(1),\ y_{\beta i}(2), \dots y_{\beta i}(N)]^T,$$

with

$$y_{\beta i}(k) = \sum_{j=0}^{k} h_j^{\beta_i} u_c(k-j),$$

where $h_j^{\beta_i}$ corresponds to impulse response of the basis β_i.

Notice that, in a noise-free case, i.e., $v(k) = 0$, it is possible to find a controller $C(\theta) = C(\theta^0)$, such that $\overline{e}_m(\theta, k) = 0$. However, in any practical setting we have $\delta_p > 0$ and then, the Set Membership approach is suitable to find a set containing all the controllers $C(\theta)$ that are compatible with data, noise bound δ_p and *a priori* information on the reference model M and the controller structure, defined by the basis set.

Under the definition of the FPS and previous assumptions, the next theorem is stated.

Theorem 2. *Given a data set $\mathcal{D}$, a reference model $M(q^{-1})$ and a set of basis functions $\{\beta_1(q^{-1}),\ \beta_2(q^{-1}),\ \dots,\ \beta_{m_{max}}(q^{-1})\}$. If $\delta_p \geq \delta_p^{min}$, for δ_p^{min} the solution to the convex optimization problem*

$$\begin{aligned} \delta_p^{min} = \quad & \min\nolimits_{\theta\in\Theta} \delta \\ & s.t. \\ & \|Y_c - \Phi\theta\|_{\ell_p} \leq \delta \\ & \delta \geq 0 \end{aligned} \tag{17}$$

Then, $FPS \neq \emptyset$.

Proof:

Note that θ^*, the argument minimizing (17), guarantees,

$$\|Y_c - \Phi\theta^*\|_{\ell_p} \leq \delta_p$$

Then, $\theta^* \in FPS$.

The previous theorem gives a tool to determine if a set of *a priori* hypotheses is compatible with the available data. For example, it allows to evaluate if a reference model is achievable with the selected basis functions with acceptable error.

3.2. Finding a sparse controller

The FPS defined in the previous subsection does not take into account assumption 3, that is, the feasible controllers can have any cardinality. Under Assumption 3, a feasible parameters set can be defined, where controllers have the structure of $C^o(\theta^o)$, i.e.,

Definition 2. *Feasible Sparse Parameters Set (FSPS):*

$$\begin{aligned} FSPS = \big\{&\theta \in \Theta : supp(\theta) = supp(\theta^0), \\ &\|Y_c - \Phi\theta\|_{\ell_p} \leq \delta_p\Big\} \end{aligned} \tag{18}$$

However, the support of the optimal controller is unknown. The sparsest controller, compatible with hypotheses and experimental data might be found solving the following optimization problem:

$$\begin{aligned} \theta^0 = &\arg\min_{\theta\in\Theta} \|\theta\|_0 \\ &s.t. \\ &\|Y_c - \Phi\theta\|_{\ell_p} \leq \delta_p \end{aligned} \tag{19}$$

In fact, maximizing the sparsity of a vector corresponds to minimizing its l_0 quasi-norm. However, the ℓ_0 quasi-norm is a non-convex function and its minimization is, in general, an NP-hard problem. In an identification framework, convex relaxations, see e.g. [26, 27, 28], and greedy algorithms, see e.g. [29], are the main approaches to deal with this problem.

Instead of minimizing the support of the controller, a different approach is to limit the complexity of the controllers set to a fixed number of basis functions m_θ. This leads to the next limited complexity feasible parameters set:

Definition 3. *Limited complexity Feasible Parameter Set (FPS):*

$$FPS(m_\theta) = \Big\{\theta \in \Theta : \|Y_c - \Phi\theta\|_{\ell_p} \leq \delta_p \ \wedge \ \|\theta\|_0 = m_\theta\Big\} \tag{20}$$

Note that $FPS(m_\theta)$ is the union of $\binom{m_{max}}{m_\theta}$ subsets with the same cardinality and, to fully characterize the set, it is necessary to verify the feasibility condition established in Theorem 2 for each sub-set of bases guaranteeing $||\theta||_0 = m_\theta$, this is a combinatorial problem intractable for large sets of basis functions.

Instead of verifying the feasibility of each sub-set in $FPS(m_\theta)$, in this work we propose an "smart" selection of active basis functions, following an approach similar to [30], where a greedy algorithm is proposed in the context of system identification for fault detection.

The proposed limited-complexity controller estimation algorithm has two steps. First, the support of a particular controllers set is estimated and then, an interpolatory solution is provided.

Consider the optimization problem

$$\begin{aligned} \theta^1 = &\arg\min_{\theta\in\Theta} \|\theta\|_1 \\ &s.t. \\ &\|Y_c - \Phi\theta\|_{\ell_p} \leq \delta_p \end{aligned} \tag{21}$$

where δ_p guarantees a non-empty FPS, according to *Theorem* 2. Recall that the ℓ_1 norm is the convex envelope of the ℓ_0 quasi-norm, and its minimization yields a sparse vector θ^1 [26, 27, 28].

From the optimal solution θ^1 obtained in (21), the following definition provides an ordered set of bases.

Definition 4. *Let*

$$r(\theta^1) \doteq \left\{i_1, ...i_{m_{max}} : \left|\theta^1_{i1}\right| \geq ... \geq \left|\theta^1_{im_{max}}\right|\right\}.$$

For any $m_\theta \in \{1, 2, \ldots, m_{max}\}$, the support of any controller of complexity m_θ is

$$\lambda(m_\theta) = \{i_1, i_2, \ldots, i_{m_\theta}\} \tag{22}$$

Note that $r(\theta^1)$ is the set of basis indexes, sorted by the amplitude of the elements of θ^1.

From the previous ordered set of bases, the Feasible Sparse Parameters Set is defined as:

Definition 5. *Feasible Sparse Parameters Set (FSPS) of complexity m_θ:*

$$\begin{aligned} FSPS(m_\theta) = \Big\{\theta \in \Theta : \|Y_c - \Phi\theta\|_{\ell_p} \leq \delta_s \ \ \wedge \\ \theta_i = 0, \forall i \notin \lambda(m_\theta)\} \end{aligned} \tag{23}$$

Lemma 1. *If $\delta_s \geq \delta_s^{min}(m_\theta)$, for $\delta_s^{min}(m_\theta)$ the solution to the convex optimization problem*

$$\begin{aligned} \delta_s^{min}(m_\theta) = \quad & \min\nolimits_{\theta\in\Theta} \delta \\ & s.t. \\ & \|Y_c - \Phi\theta\|_{\ell_p} \leq \delta \\ & \delta \geq 0 \\ & \theta_i = 0, \forall i \notin \lambda(m_\theta) \end{aligned} \tag{24}$$

Then, $FSPS(n_\theta) \neq \emptyset$.

Corollary 3. *Let $\epsilon(m_\theta) = \delta_s^{min}(m_\theta) - \delta_p^{min}$, then $\epsilon(m_\theta) \geq 0$.*

Moreover, if $\epsilon(m_\theta) = 0$, then $FSPS(m_\theta) \subset FPS$.

The previous results follow from the fact that, for any δ, the feasible set in optimization problem (17) contains the feasible set in problem (24).

For any m_θ, such that $\epsilon(m_\theta) = 0$, the selected subset of basis functions guarantee the same accuracy explaining the available data than the full basis set. It follows that the behavior of $\epsilon(m_\theta)$ is an indicator of the trade-off between accuracy and sparsity, allowing the designer to select a proper controller complexity m_θ.

Once the support of the controller is defined, the final step in the controller tuning procedure is to select a vector of parameters $\hat{\theta}$ belonging to $FSPS(m_\theta)$, guaranteeing an small closed-loop error. Two solutions are proposed:

- The following optimization problem allows to identify an interpolatory estimate, minimizing the controller output error on the available data set:

$$\begin{aligned}\hat{\theta}_I \quad = \quad & \arg \min_{\theta \in \Theta} \|Y_c - \Phi(t)\theta\|_{\ell_p} \\ & s.t. \\ & \theta_i = 0, \forall i \notin \lambda(m_\theta)\end{aligned} \tag{25}$$

- A central estimate, given by the Chebyshev center of the $FSPS(m_\theta)$, can also be employed, minimizing the worst-case error in the parameter space, but increasing the computational complexity of the tuning process.

$$\begin{aligned}\hat{\theta}_C \quad = \quad & \arg \min_{\theta} \max_{\theta' \in FSPS(m_\theta)} \|\theta - \theta'\|_{\ell_q} \\ & s.t. \\ & \theta_i = 0, \forall i \notin \lambda(m_\theta)\end{aligned} \tag{26}$$

Now we are in a position to propose a limited-complexity controller estimation algorithm as follows.

Algorithm 1. *Sparse Set Membership tuning algorithm (SSMT)*

1. *Collect a data set $\mathcal{D}$ performing an experiment starting with the plant initially at rest.* Note that the assumptions on noise do not require open-loop operation.
2. *Select a proper reference model M and basis functions set $\{\beta_1, \beta_2, \ldots, \beta_{m_{max}}\}$.*
3. *Obtain a lower noise bound δ_p (see Theorem 2).*
4. *Solve the optimization problem*

$$\begin{aligned}\theta^1 = & \arg \min_{\theta \in \Theta} \|\theta\|_1 \\ & s.t. \\ & \|Y_c - \Phi\theta\|_{\ell_p} \leq \delta_p\end{aligned} \tag{27}$$

   and construct vector $r(\theta^1)$ as in Definition *4.*
5. *Select an sparsity error tolerance bound ϵ_{max} and perform the following iteration:*

for $j = 1 : m_{max} - 1$

$$\begin{aligned} m_\theta = & \ m_{max} - j \\ \theta(j) = & \ \arg\min_{\theta\in\Theta} \|Y_c - \Phi\theta\|_{\ell_p} \\ & \ s.t. \\ & \ \theta_i = 0, \forall i \notin \lambda(m_\theta) \end{aligned}$$

$$\text{if} \quad \|Y_c - \Phi\theta(j)\|_{\ell_p} \leq \delta_p + \epsilon_{max}$$

$$\hat{\theta} = \theta(j)$$

else

break

end

After, arranging the elements of θ^1 in decreasing amplitude order, in step 5 the bases with smaller coefficients are removed, one by one, until the given threshold ϵ_{max} for the output error increment is reached.

Remark 2. *ϵ_{max} can be tuned to suitably manage the trade-off between accuracy and sparsity in step 5, since large values of ϵ_{max} lead to large sparsity, that is, $m_\theta \ll m_{max}$.*

Under suitable conditions on Φ and θ^1, the results in [31] guarantee that $\hat{\theta}$, derived by Algorithm 1, is maximally sparse with the same support as θ^o.

4. Numerical Case Study

In this section, the proposed approach is evaluated on simulated data, generated by a flexible-transmission model, comparing its performance against a sparse-CbT algorithm, proposed in [20]. As far as the authors are aware, it is the only reported method that uses a sparse approach to solve Problem 1.

Consider the flexible transmission system introduced as a benchmark for digital control design in [32]. The plant transfer function is:

$$P(q^{-1}) = \frac{0.28261q^{-3} + 0.50666q^{-4}}{1 - 0.418q^{-1} + 1.589q^{-2} - 1.316q^{-3} + 0.886q^{-4}}$$

The control objective is given in terms of model-reference specifications, given by:

$$M(q^{-1}) = \frac{(1-\alpha)^2 q^{-3}}{(1-\alpha q^{-1})^2}$$

where α indicates the location of the poles defining the desired loop speed and bandwidth. As in previous works with this benchmark, it is used $Ts = 0.05$ s and $\alpha = 0.6$. The basis functions parametrizing the controller are:

$$\beta_i(q^{-1}) = \frac{q^{1-i}}{1-q^{-1}}, i = 1, 2, ..., m_{max}.$$

It is assumed $m_{max} = 12$.

First, a data set of $N = 200$ samples is generated with the system operating in open loop, input $u(k)$ is generated as a sequence of i.i.d. samples of a Gaussian distribution with zero mean and variance 4. The plant output $w(k)$ is affected by additive output noise $v(k)$ generated as i.i.d. samples of a uniform distribution with zero mean an a Signal to Noise Ratio (SNR) of $20dB$. The SNR is calculated as

$$SNR = 10log\frac{\sum_{t=1}^{N} y(t)^2}{\sum_{t=1}^{N} v(t)^2}.$$

Then, Theorem 2 is employed to determine lower bounds on the output noise norm δ_p^{min}. $p = 2$ is selected as signal norm and different reference models (i.e different α values) are tested for several m_{max} values. In this case, all the m_{max} basis functions are employed. Results are shown in Figure 3. Note that, for fast reference models (α small), the noise bound is high even for a large set of basis. On the other side, for $\alpha = 0.6$, $\delta_2^{min} \approx 2.8$ can be selected as a validated bound.

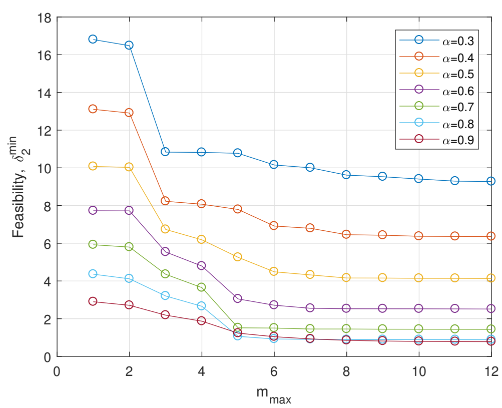

Figure 3: Lower noise bounds for different reference models and controller complexities (m_{max}).

The next step is to execute step 5 in *Algorithm 1*. Different ϵ_{max} values are considered, in order to highlight the trade-off between accuracy and sparsity. Results are reported in Figure 4. It is shown that increasing ϵ_{max} leads to higher sparsity. When $\epsilon_{max} = 0$, i.e., $FSPS \in FPS$, 10 bases are selected by the algorithm, while $\epsilon_{max} = 0.8$ offers an acceptable trade-off to obtain a "low-complexity" controller. In this case the estimated sparse-controller is:

$$C(\hat{\theta}) = \frac{0.1006 - 0.0308q^{-1} + 0.0625q^{-5}}{1 - q^{-1}}$$

The complexity is $m_\theta = 3$, and the bases selected by the algorithm are $\beta_i(q^{-1}) = \frac{q^{1-i}}{1-q^{-1}}, i = 1, 2, 6$.

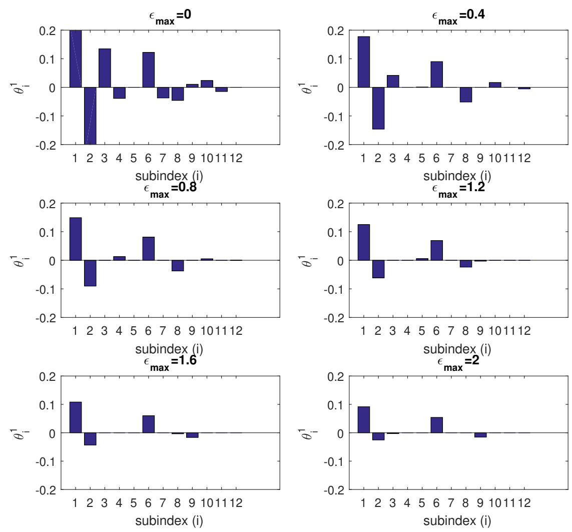

Figure 4: Sparsity results for different ϵ_{max} values.

Finally, the solution proposed in [20] is applied to the available data set. The sparse CbT algorithm relies on ℓ_1 regularization. The cost function to be minimized is:

$$\hat{\theta}^{(j)} = \arg\min_{\theta} \left[J(\theta) + \lambda \left\| W^{(j)}\theta \right\|_1 \right] \tag{28}$$

where $J(\theta)$ is the 2-norm of the correlation signal between a model error and a properly selected instrumental variable, W is a weighing diagonal matrix and j is a counter index.

Parameters $l = 30$ (instrumental variable length), $\lambda_{max} = 0.15$ (upper bound on regularization weight), $j_{max} = 10$ (maximum number of iterations), and $\epsilon = 0.01$ (lower bound of weights in matrix W) are selected, according to the criteria given in [20]. Initially, $\lambda = 0.01$ is employed but the algorithm does not converge (i.e. when $j = j_{max}$, $\|\theta\|_0 > m_\theta$), results are reported in Table 1. A second test with $\lambda = 0.12$ is performed and the results are reported in table 2. It can be seen that the algorithm converges after 3 iterations.

Table 1: SCbT convergence results for $\lambda = 0.01$

iteration	1	2	3	4	5	6	7	8	9	10
$\|\theta\|_0$	12	12	10	8	8	8	8	7	7	7

Table 2: SCbT convergence results for $\lambda = 0.12$

iteration	1	2	3	4	5	6	7	8	9	10
$\|\theta\|_0$	12	5	3	3	3	3	3	3	3	3

The sparse-controller obtained is

$$C(\hat{\theta}) = \frac{0.0951 - 0.0235q^{-1} + 0.0643q^{-5}}{1 - q^{-1}}$$

Note that the bases selected by the *SCbT* algorithm coincide with those identified by the SSMT algorithm.

4.1. Monte-Carlo test

A final test is performed, where a Monte-Carlo experiment allows to evaluate the behavior of the algorithm for 100 realizations of noise $v(t)$, leading to 100 data sets $\mathcal{D}$, all of them maintaining a $SNR \approx 20dB$.

For each data set and method, a controller has been tuned and tested on the model of the plant P. Four cases are addressed, $N = 300$, $N = 200$, $N = 100$ and $N = 70$. The controllers have been tuned as described in the previous example. In the *SCbT* method $l = 30$ or $l = 20$ is used according to data length N.

For each data set size N, the performance of the estimated controllers is measured via simulation. The quality of the resulting control action is measured as the maximum error (M_e) and the root mean square error ($RMSE$) of the closed-loop step response with respect to the reference model. The average and worst-case results are shown in Table 3.

As can be seen in Table 3, for the case of $N = 300$ the results are comparable for both methods. But, for $N = 200$, the RMSE almost doubles for the *SCbT* method. For the cases $N = 100$ and $N = 50$, both performance measures for the controllers obtained via $SSMT$ method keep constant, while those achieved via *SCbT* algorithm increase. It is highlighted that in all cases the $SSMT$ method proposed in this work shown better results than the *SCbT* method.

Figure 5 shows the percentage of experiments (noise realizations) for which, each algorithm properly estimated the optimal basis functions $\beta_i(q^{-1}) = \frac{q^{1-i}}{1-q^{-1}}, i = 1, 2, 6.$. For $N = 300$ both methods are able to recover the correct basis. However, for smaller data sets, the *SCbT* method is not always able to find the correct bases to construct the sparse-controller, even when the number of iterations $j_{max} = 15$ is selected. For $N = 70$, correct bases are obtained in less than 50% of the experiments for the *SCbT* method. This was expected because the condition $\ell/N \ll 1$ is not fulfilled, such condition is required for the correlation method (see details in []). Moreover, in some cases it produces unstable closed-loops. However, the performances measures reported in Table 3 consider only stable controllers.

Table 3: Results for Monte-Carlo experiment

Case	Measure	SSMT	SCbT
N=300,	$\overline{RMSE}/max(RMSE)$	0.0048/0.0051	0.0068/0.0093
l=30	$\overline{M_e}/max(M_e)$	0.1996/0.2025	0.2437/0.2770
N=200,	$\overline{RMSE}/max(RMSE)$	0.0050/0.0051	0.0100/0.0102
l=30	$\overline{M_e}/max(M_e)$	0.1941/0.2053	0.2641/0.2745
N=100,	$\overline{RMSE}/max(RMSE)$	0.0041/0.044	0.0235/0.1153
l=30	$\overline{M_e}/max(M_e)$	0.1914/0.1950	0.3650/0.5602
N=70,	$\overline{RMSE}/max(RMSE)$	0.0048/0.0051	0.0812/0.1133
l=20	$\overline{M_e}/max(M_e)$	0.1914/0.1950	0.2599/0.2776

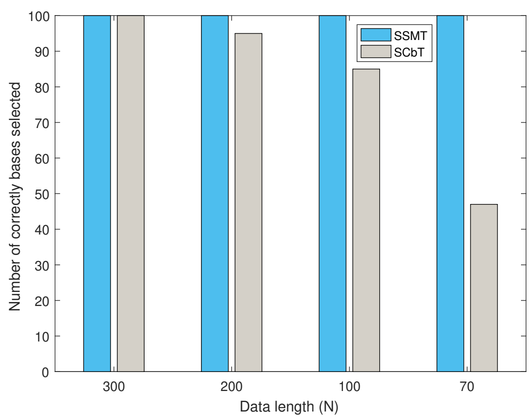

Figure 5: Results for bases selection.

5. Conclusion

We have developed a new Data-Driven approach to design limited-complexity controllers for linear systems using Set Membership techniques and sparse identification methods. An algorithm has been proposed in order to solve the sparse-controller tuning problem, supported by feasibility theorems that provide a single parameter to adjust the complexity-accuracy trade-off. The algorithm proposed here avoids solving big combinatorial problems that arise when the dimension of the vector parameterizing the candidate controllers is large and the number of desired parameters is much lower. A benchmark flexible transmission model is employed to illustrate the performance of the proposed methodology (SSMT), in comparison to the sparse correlation based tuning approach (SCbT). It is found that both approaches offer similar performance when the size of the data set is much larger than the dimension of the controller parameters vector. Notwithstanding, the SCbT controllers are strongly affected when data set size is reduced, while the SSMT controllers exhibit good performances even when the controller parameters are estimated from reduced data sets. Moreover, SSMT method has just one parameter to be adjusting and it has a direct interpretation as modelling error bound, simplifying the tuning procedure.

Acknowledgment

The work of F. Valderrama was financed by a doctoral scholarship from Gobernación de Boyacá, Colombia. (call 733 Colciencias).

Appendix A. Proof of Theorem 1

Proof: From assumption 2, it is known that

$$M = \frac{C^o P}{1 + C^o P}$$

and

$$1 - M = \frac{1}{1 + C^o P}.$$

Then, the model matching error system can be expressed as:

$$\begin{aligned} E_m(\theta) &= M - \frac{C(\theta)P}{1 + C(\theta)P} \\ &= \frac{C^o P}{1 + C^o P} - \frac{C(\theta)P}{1 + C(\theta)P} \end{aligned}$$

After some algebra we have,

$$\begin{aligned} &= \frac{C^o P - C(\theta)P}{(1 + C^o P)(1 + C(\theta)P)} \\ &= \frac{C^o P - C(\theta)P}{(1 + C^o P)(1 + (1 + \Delta(\theta))C^o P)} \\ &= \frac{C^o P - C(\theta)P}{(1 + C^o P)^2(1 + \frac{C^o \Delta(\theta) P}{1 + C^o P})} \\ &= \frac{1}{1 + M\Delta(\theta)} \left(\frac{C^o P}{(1 + C^o P)^2} - \frac{C(\theta)P}{(1 + C^o P)^2} \right) \\ &= \frac{1}{1 + M\Delta(\theta)} \big(M(1 - M) - (1 - M)^2 C(\theta)P \big) \end{aligned}$$

arriving to the stated modelling error system.

References

[1] K. Van Heusden, A. Karimi, T. Söderström, On identification methods for direct data-driven controller tuning, International Journal of Adaptive Control and Signal Processing 25 (5) (2011) 448–465.

[2] S. Formentin, K. van Heusden, A. Karimi, A comparison of model-based and data-driven controller tuning, International Journal of Adaptive Control and Signal Processing 28 (10) (2014) 882–897. doi:10.1002/acs.2415.

[3] M. Campi, A. Lecchini, S. Savaresi, Virtual reference feedback tuning: a direct method for the design of feedback controllers, Automatica 38 (8) (2002) 1337 – 1346. doi:https://doi.org/10.1016/S0005-1098(02)00032-8.

[4] M. C. Campi, S. M. Savaresi, Direct nonlinear control design: the virtual reference feedback tuning (VRFT) approach, IEEE Transactions on Automatic Control 51 (1) (2006) 14–27.

[5] L. Campestrini, D. Eckhard, L. A. Chía, E. Boeira, Unbiased MIMO VRFT with application to process control, J. Process Control 39 (2016) 35–49. doi:10.1016/j.jprocont.2015.12.010.

[6] E. Hedrea, M. Radac, R. Precup, Virtual reference feedback tuning for position control of a twin rotor aerodynamic system, in: 2016 IEEE 11th International Symposium on Applied Computational Intelligence and Informatics (SACI), 2016, pp. 57–62. doi:10.1109/SACI.2016.7507431.

[7] R. Roman, M. Radac, R. Precup, Multi-input–multi-output system experimental validation of model-free control and virtual reference feedback tuning techniques, IET Control Theory Applications 10 (12) (2016) 1395–1403. doi:10.1049/iet-cta.2016.0028.

[8] Y. Kansha, Y. Hashimoto, M.-S. Chiu, New results on VRFT design of PID controller, Chemical Engineering Research and Design 86 (8) (2008) 925–931.

[9] A. Karimi, K. van Heusden, D. Bonvin, Non-iterative data-driven controller tuning using the correlation approach, in: 2007 European Control Conference (ECC), 2007, pp. 5189–5195. doi:10.23919/ECC.2007.7068802.

[10] K. van Heusden, A. Karimi, D. Bonvin, Data-driven model reference control with asymptotically guaranteed stability, International Journal of Adaptive Control and Signal Processing 25 (4) (2011) 331–351. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/acs.1212, doi:10.1002/acs.1212.

[11] M. Milanese, C. Novara, Set membership identification of nonlinear systems, Automatica 40/6 (2004) 957–975.

[12] M. Milanese, M. Taragna, $\mathcal{H}_\infty$ set membership identification: A survey, Automatica 41 (12) (2005) 2019–2032.

[13] M. Milanese, F. Ruiz, M. Taragna, Direct data-driven filter design for uncertain lti systems with bounded noise, Automatica 46 (11) (2010) 1773–1784.

[14] F. Ruiz, C. Novara, M. Milanese, Direct design from data of optimal filters for LPV systems, Systems & Control Letters 59 (1) (2010) 1 – 8.

[15] C. Novara, F. Ruiz, M. Milanese, Direct filtering: A new approach to optimal filter design for nonlinear systems, Automatic Control, IEEE Transactions on 58 (1) (2013) 86–99.

[16] F. Valderrama, F. Ruiz, Controller design from data under UBB noise, Proc. Am. Control Conf. (2014) 5103–5108doi:10.1109/ACC.2014.6858626.

[17] V. Cerone, D. Regruto, M. Abuabiah, Direct data-driven control design through set-membership errors-in-variables identification techniques, in: Proc. Am. Control Conf., 2017, pp. 388–393. doi:10.23919/ACC.2017.7962984.

[18] S. Formentin, C. Novara, S. M. Savaresi, M. Milanese, Data-driven inversion-based control of nonlinear systems, IFAC-PapersOnLine 48 (28) (2015) 1343 – 1348, 17th IFAC Symposium on System Identification SYSID 2015. doi:https://doi.org/10.1016/j.ifacol.2015.12.318.

[19] M. Tanaskovic, L. Fagiano, C. Novara, M. Morari, On-line direct control design for nonlinear systems, IFAC-PapersOnLine 48 (28) (2015) 144 – 149, 17th IFAC Symposium on System Identification SYSID 2015. doi:https://doi.org/10.1016/j.ifacol.2015.12.115.

[20] S. Formentin, A. Karimi, Direct data-driven design of sparse controllers, in: 2013 American Control Conference, 2013, pp. 3099–3104. doi:10.1109/ACC.2013.6580307.

[21] O. Goro, B. Anderson, Model Reduction for Control System Design, Springer-Verlag London, 2001.

[22] B. D. O. Anderson, Y. Liu, Controller reduction: concepts and approaches, IEEE Transactions on Automatic Control 34 (8) (1989) 802–812. doi:10.1109/9.29422.

[23] A. Karimi, G. Galdos, Fixed-order h-infinity controller design for nonparametric models by convex optimization, Automatica 46 (8) (2010) 1388 – 1394. doi:https://doi.org/10.1016/j.automatica.2010.05.019.

[24] S. Skogestad, I. Postlethwaite, Multivariable feedback control: analysis and design, John Wiley, 2005.

[25] K. Zhou, J. C. Doyle, Essentials of Robust Control, Prentice Hall International, 1998.

[26] J. A. Tropp, Just relax: convex programming methods for identifying sparse signals in noise, IEEE Transactions on Information Theory 52 (3) (2006) 1030–1051. doi:10.1109/TIT.2005.864420.

[27] D. L. Donoho, M. Elad, V. N. Temlyakov, Stable recovery of sparse overcomplete representations in the presence of noise, IEEE Transactions on Information Theory 52 (1) (2006) 6–18. doi:10.1109/TIT.2005.860430.

[28] J. J. Fuchs, Recovery of exact sparse representations in the presence of bounded noise, IEEE Transactions on Information Theory 51 (10) (2005) 3601–3608. doi:10.1109/TIT.2005.855614.

[29] J. A. Tropp, Greed is good: algorithmic results for sparse approximation, IEEE Transactions on Information Theory 50 (10) (2004) 2231–2242. doi:10.1109/TIT.2004.834793.

[30] C. Novara, Sparse set membership identification of nonlinear functions and application to fault detection, International Journal of Adaptive Control and Signal Processing 30 (2) (2016) 206–223. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/acs.2539, doi:10.1002/acs.2539.

[31] C. Novara, Sparse identification of nonlinear functions and parametric set membership optimality analysis, IEEE Transactions on Automatic Control 57 (12) (2012) 3236–3241. doi:10.1109/TAC.2012.2202051.

[32] I. Landau, D. Rey, A. Karimi, A. Voda, A. Franco, A flexible transmission system as a benchmark for robust digital control*, European Journal of Control 1 (2) (1995) 77 – 96. doi:https://doi.org/10.1016/S0947-3580(95)70011-5.