

Asymptotically exact direct data-driven multivariable controller tuning

Simone Formentin^{*}, Andrea Bisoffi^{**}, Tom Oomen^{***}

^{*} *Dipartimento di Elettronica, Informazione e Bioingegneria,
Politecnico di Milano, via G. Ponzio, 34/5 - 20133 Milano, Italy.
E-mail to: simone.formentin@polimi.it.*

^{**} *Department of Industrial Engineering, University of Trento,
Via Sommarive, 9 - 38123 Trento, Italy.*

^{***} *Control Systems Technology Group, Department of Mechanical
Engineering, Eindhoven University of Technology, Eindhoven,
The Netherlands.*

Abstract: In this paper, a data-driven controller design method for multivariable systems is introduced and analyzed. The proposed technique is direct, as it is entirely based on experimental data and does not rely on a physical description of the system, and non-iterative, as it does not require controller adjustments based on additional experiments. Compared to the state-of-the-art non-iterative technique, *i.e.* MIMO VRFT, the proposed approach is asymptotically exact, in that it guarantees that the desired closed-loop dynamics is matched when the number of data tends to infinity. The performance of the proposed approach is illustrated and compared with MIMO VRFT on a benchmark simulation example.

Keywords: direct data-driven control; multivariable systems; convex optimization.

1. INTRODUCTION

Model based control design is a systematic approach to deal with interaction in multivariable systems (see Skogestad and Postlethwaite (2005); Maciejowski (1989)). However, modeling is expensive and it is estimated that about 75% of the cost of complex control applications is devoted to modeling, see Gevers (2005). This is particularly the case for multivariable systems. In addition, when the resulting controller structure is fixed, which is often the case (*e.g.*, multivariable PIDs), then additional issues arise, including fixed structure synthesis or model reduction, see Bunse-Gerstner et al. (2010).

The modeling of the system can be avoided, by directly optimizing the controller parameters using experimental data. The first papers on data-to-controller design proposed self-tuning regulation (STR) and model-reference adaptive control (MRAC) (see Astrom and Wittenmark (1994) for a complete overview).

More recently, two iterative methods have been proposed to design reliable controllers off-line with a small number of closed-loop experiments on the plant.

The first method is the Iterative Feedback Tuning (IFT) approach (see Hjalmarsson et al. (2002)), where an unbiased estimate of the gradient is provided entirely from input/output (I/O) data collected on the actual closed-loop system.

The second technique is known as Iterative Correlation-based Tuning (ICbT) and has been introduced in Mišković et al. (2003) for single-input/single-output (SISO) plants. A MIMO extension is discussed in Mišković et al. (2007), where it has been shown that the correlation-based controller tuning provides better tracking performance in fewer experiments than IFT.

In Formentin and Savaresi (2011) and Formentin et al. (2012), a non-iterative method has been proposed, which provides a multivariable controller based on the well known Virtual Reference Feedback Tuning (VRFT) rationale (see Campi et al. (2002); Savaresi and Guardabassi (1998)). The method relies on a single set of open-loop I/O data collected on a stable MIMO LTI square plant. Initial results in this direction also appeared in Nakamoto (2003), but this work was a preliminary attempt towards a complete strategy, as the proposed method was a straightforward extension of the existing SISO VRFT, whereas important issues as undermodeling and treatment of noise were not discussed.

In this work, a different approach to the problem is presented. Instead of the VRFT rationale in Formentin et al. (2012), a data-based reformulation of the closed-loop model matching cost function is presented and proven to be equivalent to the original problem. Then, it is shown that, by feeding the system with a specific sequence of persistently exciting signals, the reformulated problem can be solved using only the collected data and convex optimization tools, without modeling the system. The results exploit recent results from Oomen et al. (2014), where a data-driven approach to estimate the H_∞ norm of multivariable systems is investigated. Here, it is instead used for tuning MIMO controller parameters.

To deal with output noise and obtain a consistent estimate of the optimal controller, the above methodology is endowed of an instrumental variable-based correction inspired to Stoica and Jansson (2000). It comes out that, in its more general form, the algorithm can be based on simple least squares estimation.

The remainder of the paper is organized as follows. Section 2 formally states the problem and defines the math-

emathical framework. The MIMO VRFT method is briefly recalled in Section 3, where its limitations are also analyzed. The method proposed in this paper is introduced and discussed in the noiseless case in Section 4 and for noisy data in Section 5. In Section 6, the performance of the proposed method is illustrated and compared to MIMO VRFT and other state of the art techniques on a benchmark example. Some concluding remarks end the paper.

2. PROBLEM FORMULATION

Consider the *unknown* LTI discrete-time multivariable $n \times n$ stable plant $G(q^{-1})$, where q^{-1} denotes the backward shift operator. The above system is such that, provided the n -dimensional input vector $u(t) = [u_1(t), \dots, u_n(t)]^T$, the n -dimensional output $y(t) = [y_1(t), \dots, y_n(t)]^T$ is given by

$$y(t) = G(q^{-1})u(t). \quad (1)$$

Consider that a linear, fixed-order controller class $C(q^{-1}, \theta)$, parameterized through θ , is given. From now on, the dependence on q^{-1} will be sometimes omitted for brevity.

Define the \mathcal{L}_2 -norm of a generic discrete-time multivariable system $X(q^{-1})$ as

$$\|X\|_2 = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \text{tr} [X(e^{j\omega})X^H(e^{j\omega})] d\omega \right)^{1/2},$$

where the operator $\text{tr}(\cdot)$ denotes the trace of a matrix and the superscript H indicates the hermitian conjugate of a complex matrix. The control problem considered in this work is a classical model-reference control problem, that is, the problem of designing a controller in the class $C(q^{-1}, \theta)$ for which the *output* complementary sensitivity function matches a user-defined stable strictly proper reference model $M(q^{-1})$. More formally, the problem can be stated as follows.

Problem 1. Find

$$\hat{\theta}_o = \arg \min_{\theta} J_{MR}(\theta), \quad (2a)$$

$$J_{MR}(\theta) = \left\| M - (I + GC(\theta))^{-1} GC(\theta) \right\|_2^2, \quad (2b)$$

where I is the $n \times n$ identity matrix.

Consider then the following assumption and problem formulation.

Assumption 1. The n_c -th order control law is defined as

$$u(t) = u(t-1) + \sum_{m=0}^{n_c} B_m e(t-m), \quad (3)$$

where $e(t) = r(t) - y(t)$ represents the tracking error computed from the n -dimensional reference vector $r(t) = [r_1(t), \dots, r_n(t)]^T$ and $B_m \in \mathbb{R}^{n \times n}$ are matrices containing the controller parameters θ such that

$$\theta = [\text{vec}(B_0)^T \dots \text{vec}(B_{n_c})^T]^T, \quad (4)$$

where vec is the standard vectorization operator for a matrix.

Problem 2. Find

$$\hat{\theta}_c = \arg \min_{\theta} J(\theta), \quad (5a)$$

$$J(\theta) = \|\Delta(\theta)\|_2^2, \quad (5b)$$

$$\Delta(\theta) = M - (I - M)GC(\theta). \quad (5c)$$

Some remarks are due:

- Problem 1 is non-convex, whereas, under Assumption 1, Problem 2 is convex.
- The controller parameterization in (3) includes all PID-like control structures.
- If the desired sensitivity function $I - M$ is close to the actual sensitivity function $(I + GC(\hat{\theta}_o))^{-1}$, the criterion $J(\theta)$ is a good approximation of $J_{MR}(\theta)$ and $\hat{\theta}_c \approx \hat{\theta}_o$. Even if using Problem 2 as a reformulation of Problem 1 generally yields only an approximation of the controller achieving M , this choice allows the design procedure to be convexified. For this reason, similar problem reformulations are widely employed in identification for control, data-driven tuning and \mathcal{H}_2 model-reduction (see Hjalmarsson (2005) for an overview).
- $C(\hat{\theta}_o)$ generally does not make $J_{MR} = 0$, as the controller achieving $J_{MR} = 0$ might be of very high order and non-causal.

Consider now that an open-loop collection of Input/Output (I/O) data is available, namely $\mathcal{D}^N = \{u(t), \tilde{y}(t)\}$, $t = 1, \dots, N$, where

$$\tilde{y}(t) = y(t) + w(t) = G(q^{-1})u(t) + w(t) \quad (6)$$

is the noise-affected output and $w(t)$ can be any colored zero-mean noise *uncorrelated with the input* $u(t)$ and representing, *e.g.*, the measurement disturbances.

In standard “indirect” data-driven approaches, minimization of (5) can be achieved by identifying from data a model \hat{G} of the plant and evaluating $J(\theta)$ using \hat{G} . This approach is very sensitive to modeling errors, therefore structure and order selection need to be carried out very accurately.

In this work, Problem 2 will be solved *directly from data*, without the need to parameterize and identify a model \hat{G} of the system.

3. THE MIMO VRFT METHOD REVISITED

In this section, we briefly recall the state of the art technique for multivariable data-driven controller tuning, namely the MIMO VRFT approach in Formentin et al. (2012).

In few words, the idea proposed in Formentin et al. (2012) to solve Problem 2 without identifying $G(q^{-1})$ is to build a “virtual” closed-loop system, where the input and output signals are equal to $u(t)$ and $y(t)$ and the closed-loop transfer function is assumed to correspond to $M(q^{-1})$. From the above loop, the so-called “virtual reference” $r_V(t)$ and “virtual error” $e_V(t)$ signals can be computed off-line as

$$r_V(t) = M^{-1}(q^{-1})y(t), \quad e_V(t) = r_V(t) - y(t).$$

The control design problem is then reduced to an identification one and the optimization procedure is still convex if the controller structure is selected as in (3).

The problem to be solved is then, in the noiseless case, the following.

Problem 3. Find

$$\hat{\theta}_{VR} = \arg \min_{\theta} J_{VR}^N(\theta) \quad (7a)$$

$$J_{VR}^N(\theta) = \frac{1}{N} \sum_{t=1}^N \|u_{L_u}(t) - C(\theta)e_{L_e}(t)\|_2^2, \quad (7b)$$

where

$$\begin{aligned} u_{L_u}(t) &= L_u(q^{-1})u(t), \\ e_{L_e}(t) &= L_e(q^{-1})e_V(t) \\ &= L_e(q^{-1})(M^{-1}(q^{-1}) - I)G(q^{-1})L_y(q^{-1})u(t), \end{aligned}$$

and L_u, L_e, L_y are suitable data prefilters.

The prefilters are required in case the controller that leads the cost function $J(\theta)$ to zero is not in the controller set (see Campi et al. (2002)), as $\hat{\theta}_{VR}$ and $\hat{\theta}_c$ could not coincide. Optimal filter selection is defined by the following result.

Proposition 1. If data prefilters in (7) are selected as

$$L_u = M\Phi_u^{-1/2}, L_e = C^{-1}(\theta), L_y = C(\theta)\Phi_u^{-1/2}, \quad (8)$$

where Φ_u is the power spectral density of u and $\Phi_u^{1/2}$ is a spectral factor of Φ_u , then $\hat{\theta}_{VR}$ and $\hat{\theta}_c$ asymptotically coincide.

Proof. See Formentin et al. (2012). \square

Since (8) are θ -dependent, such filters cannot be implemented. Therefore, Formentin et al. (2012) employs $L_u = L_e = L = M$ and $L_y = I$. This choice is optimal only in case $C(\theta), M$ and G can commute (*e.g.*, for SISO systems), but it is suboptimal everywhere else. From a system-theoretic perspective, the above filter selection makes the frequency-expression of $J(\theta)$ in (5) equal to a new cost function $\tilde{J}(\theta)$

$$\begin{aligned} \tilde{J}(\theta) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{tr} [M - C(\theta)G(I - M)] \times \\ &\quad \times \Phi_u [M - C(\theta)G(I - M)]^H d\omega. \end{aligned} \quad (9)$$

Notice that (9) is also the frequency-wise convex approximation of

$$\tilde{J}_{MR}(\theta) = \left\| \left(M - (I + C(\theta)G)^{-1} C(\theta)G \right) \Phi_u^{1/2} \right\|_2^2. \quad (10)$$

In this way, $C(\theta)$ is chosen such to make the *input* - and not the *output* - complementary sensitivity function as close as possible to M .

In what follows, a procedure guaranteeing the matching of the *output* complementary sensitivity function is provided. In the new procedure, no prefilters will be required, thus not needing any suboptimal approximation.

4. MULTIVARIABLE CONTROLLER TUNING

In the last section, the deficiencies of the VRFT method for MIMO systems have been revealed. In this section, a new method is proposed to solve these deficiencies, which constitutes the main contribution of this paper.

To introduce the method, the case of noiseless data will be first addressed, namely, $w(t) = 0, \forall t$ in (6) so that the measured $\tilde{y}(t)$ coincides with $y(t) = G(q^{-1})u(t)$. Consider the following signal-based reformulation of Problem 2.

Problem 4. Find

$$\hat{\theta}_D = \arg \min_{\theta} J_D(\theta), \quad (11a)$$

$$J_D(\theta) = \|\varepsilon(t, \theta)\|_2^2, \quad (11b)$$

$$\varepsilon(t, \theta) = \Delta(q^{-1}, \theta)u(t), \quad t = 1, \dots, N. \quad (11c)$$

Next, the matching problem is investigated. The following result holds.

Theorem 1. (Asymptotic equivalence of Problems 2 and 4). Assume that $\Phi_u(\omega) = \lambda^2 I, \forall \omega, \lambda \in \mathbb{R}$. Then,

$$\lim_{N \rightarrow \infty} \hat{\theta}_D = \hat{\theta}_c.$$

Proof. By definition of \mathcal{L}_2 norm of a MIMO system, (5b) can be rewritten as

$$J(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{tr} [\Delta(e^{-j\omega}, \theta)] [\Delta(e^{-j\omega}, \theta)]^H d\omega. \quad (12)$$

For Parseval's theorem, (11b) is such that

$$\lim_{N \rightarrow \infty} J_D(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{tr} [\Delta(e^{-j\omega}, \theta)] \Phi_u(\omega) [\Delta(e^{-j\omega}, \theta)]^H d\omega. \quad (13)$$

Under the required assumption on $\Phi_u(\omega)$, it holds that $\lim_{N \rightarrow \infty} J_D(\theta) = \lambda^2 J(\theta)$, then the values of θ minimizing $\lim_{N \rightarrow \infty} J_D(\theta)$ and $J(\theta)$ coincide for any λ , which is the thesis. \square

Although in Problem 4 the knowledge of a model of G is still required to compute $\varepsilon(t, \theta)$, in the new formulation of this work the design problem can be solved using data only and bypassing the identification of G .

As a matter of fact, notice that the matching signal $\varepsilon(t, \theta)$ can be computed as

$$\begin{aligned} \varepsilon(t, \theta) &= \Delta(q^{-1}, \theta)u(t) \\ &= M(q^{-1})u(t) - (I - M(q^{-1}))G(q^{-1})C(q^{-1}, \theta)u(t) \\ &= M(q^{-1})u(t) - (I - M(q^{-1}))v(t, \theta), \end{aligned} \quad (14)$$

where the i -th element of $v(t)$ is expressed by

$$v_i(t, \theta) = \sum_{j=1}^n \sum_{k=1}^n G_{ij} C_{jk}(\theta) u_k(t) \quad (15)$$

and $u_k(t)$ is the k -th entry of $u(t)$. Since G_{ij} and C_{jk} are scalar systems, they can commute and, therefore, (15) can equivalently be rewritten as

$$v_i(t, \theta) = \sum_{j=1}^n \sum_{k=1}^n C_{jk}(\theta) G_{ij} u_k(t) = \sum_{j=1}^n \sum_{k=1}^n C_{jk}(\theta) y_{ijk}(t), \quad (16)$$

with obvious definition for $y_{ijk}(t)$. Notice that, using (16), (14) is no longer expressed as a function of the plant model. A similar approach has been proposed in Oomen et al. (2014) to directly estimate from data the \mathcal{H}_{∞} norm of a dynamical system.

From an experimental point of view, $y_{ijk}(t)$ can be obtained from $n \times n$ experiments as follows.

Define n sequences of input signals $u_1(t), \dots, u_n(t)$ with $t = 1, \dots, N$ such that (i) each one of them may coincide with $u_k(t)$ in (15) and (16) (ii) when stacked into a vector $[u_1(t) \dots u_n(t)]^T$, such a vector actually represents the vector of signals $u(t)$ in (14). For consistency, it will be assumed from now on that the input $u(t)$ used in the identification experiment satisfies the assumption of Theorem 1.

Notice that, when the input sequence k is positioned on the input channel j , *i.e.*, the input is

$$u(t) = [\underbrace{0}_{\text{position 1}} \dots 0 \underbrace{u_k(t)}_{\text{position } j} 0 \dots \underbrace{0}_{\text{position } n}]^T,$$

the output corresponds to

$$y_{jk}(t) = G(q^{-1})u(t) = G^{(j)}(q^{-1})u_k(t) \quad (17)$$

where $G^{(j)}(q^{-1})$ denotes the j -th column of transfer function $G(q^{-1})$. The generic component i of this $y_{jk}(t)$ is exactly the term $y_{ijk}(t)$ in (16). It is clear then that $n \times n$ open-loop experiments are needed to get all the $y_{jk}(t)$ vectors in (17).

Finally, it should be here remarked that the n sequences of input signals $u_k(t)$, $k = 1, \dots, n$, $t = 1, \dots, N$, must be selected such that every channel of the real-world plant can be fed with any sequence of the set.

5. DEALING WITH NOISY DATA

In the last section, a new data-driven method was presented to determine the optimal solution to Problem 2. In this section, an approach to deal with the presence of output noise is discussed.

Suppose that $w(t) \neq 0$ in (6). It follows that all the y_{ijk} 's are noisy and the terms $G_{ij}u_k(t)$ cannot be equivalently replaced by them. In other words, the minimization of $J_D(\theta)$ would yield biased results of the estimate, as it will be shown next.

In this Section, the instrumental variable technique proposed in Stoica and Jansson (2000) will be employed to make the minima of noisy and noiseless cost criteria coincident, similarly to the procedure followed in Formentin et al. (2012).

To start with, define the estimates of the v_i 's in (16) as

$$\hat{v}_i(t, \theta) = \sum_{j=1}^n \sum_{k=1}^n C_{jk}(\theta) \tilde{y}_{ijk}(t), \quad (18)$$

using (6), which yields:

$$\hat{v}_i(t, \theta) = v_i(t, \theta) + \sum_{j=1}^n \sum_{k=1}^n C_{jk}(\theta) w_{ijk}(t) \quad (19)$$

and clearly $\hat{v}_i(t, \theta) \neq v_i(t, \theta)$. The bias term is null when $\theta = 0$, therefore the minimization of the computable matching error

$$\hat{\varepsilon}(t, \theta) = M(q^{-1})u(t) - (I - M(q^{-1}))\hat{v}(t, \theta) \quad (20)$$

would undesirably lead to a trade-off solution between 0 and $\hat{\theta}_D$.

Introduce now the extended instrumental variable $\zeta(t)$ (see Stoica and Jansson (2000)) as

$$\zeta(t) = \begin{bmatrix} u(t+l) \\ \vdots \\ u(t-l) \end{bmatrix},$$

where l is a sufficiently large integer.

Since by assumption w and u are uncorrelated, then

$$\begin{aligned} \Xi(\theta) &= \mathbb{E}[\hat{\varepsilon}(t)\zeta^T(t)] \\ &= \mathbb{E}[(M(q^{-1})u(t) - (I - M(q^{-1}))\hat{v}(t, \theta))\zeta^T(t)] \\ &= \mathbb{E}[(M(q^{-1})u(t) - (I - M(q^{-1}))v(t, \theta))\zeta^T(t)] \\ &= \mathbb{E}[\varepsilon(t)\zeta^T(t)] \end{aligned} \quad (21)$$

(see Stoica and Jansson (2000) for further details). Ideally, the optimal solution should make $\varepsilon(t) = 0$, $\forall t$, therefore one way to search for the solution $\hat{\theta}_D$ in the noisy setting is to enforce $\Xi(\theta) = 0$.

From the above reasoning, it follows that

$$\text{vec} \left\{ \mathbb{E} \left[\left(M(q^{-1})u(t) - (I - M(q^{-1}))\hat{v}(t, \hat{\theta}_D) \right) \zeta^T(t) \right] \right\}$$

should be enforced to zero.

Now, using the controller expression in (3), the $\hat{v}_i(t, \theta)$'s in (18) can be rewritten as

$$\hat{v}_i(t, \theta) = \frac{1}{1 - q^{-1}} \sum_{j=1}^n \sum_{k=1}^n \sum_{m=0}^{n_c} B_m^{(jk)} \tilde{y}_{ijk}(t - m), \quad (22)$$

where $B_m^{(jk)}$ is the element in position (j, k) of B_m .

The triple sum in (22) is equivalent to

$$\underbrace{\left[\tilde{y}_{i11}^{(I)}(t) \dots \tilde{y}_{inn}^{(I)}(t) \dots \tilde{y}_{i11}^{(I)}(t - n_c) \dots \tilde{y}_{inn}^{(I)}(t - n_c) \right]}_{\tilde{Y}_i^{(I)}(t)} \theta,$$

where $y_{ijk}^{(I)}(t)$ is the discrete integral of $y_{ijk}(t)$. Then, each component of $\hat{v}(t, \theta)$, $\hat{v}_i(t, \theta)$, can be written as $\hat{v}_i(t, \theta) = \tilde{Y}_i^{(I)}(t)\theta$ and

$$\hat{v}(t, \theta) = \tilde{Y}^{(I)}\theta \quad (23)$$

where $\tilde{Y}^{(I)}$ is the stack of the $\tilde{Y}_i^{(I)}$'s.

Then, from (21) and (23), the optimal solution needs to set

$$\mathbb{E} \left[\text{vec} \left\{ \left(M(q^{-1})u(t) - (I - M(q^{-1}))\tilde{Y}^{(I)}(t)\hat{\theta}_D \right) \zeta^T(t) \right\} \right]$$

to zero.

Thanks to the properties of the Kronecker product \otimes ,

$$\begin{aligned} &\mathbb{E} \left[\text{vec} \left\{ (I - M(q^{-1}))\tilde{Y}^{(I)}(t)\hat{\theta}_D \zeta^T(t) \right\} \right] \\ &= \mathbb{E} \left[\left(\zeta(t) \otimes ((I - M(q^{-1}))\tilde{Y}^{(I)}(t)) \right) \hat{\theta}_D \right]. \end{aligned}$$

Then, by splitting $\mathbb{E}[\cdot]$ into its single terms, it can be highlighted that $\Xi(\hat{\theta}_D) = 0$ if

$$r - R\hat{\theta}_D = 0,$$

where

$$r = \mathbb{E}[\text{vec}\{u_M(t)\zeta^T(t)\}] \quad (24a)$$

$$R = \mathbb{E}[\zeta(t) \otimes ((I - M(q^{-1}))\tilde{Y}^{(I)}(t))] \quad (24b)$$

and $u_M(t) = M(q^{-1})u(t)$.

The solution $\hat{\theta}_D$ can be easily computed using least squares-like formulas and the sample versions of R and r , that is,

$$\hat{\theta}_D = \left(\hat{R}^T \hat{R} \right)^{-1} \hat{R}^T \hat{r}, \quad (25)$$

$$\hat{R} = \frac{1}{N} \sum_{t=1}^N \zeta(t) \otimes ((I - M(q^{-1}))\tilde{Y}^{(I)}(t)) \quad (26a)$$

$$\hat{r} = \frac{1}{N} \sum_{t=1}^N \text{vec}\{u_M(t)\zeta^T(t)\} \quad (26b)$$

where summations could also be made from t sufficiently large instead of from $t = 0$, in order to get rid of the effect of the initial conditions, as indicated in Stoica and Jansson (2000).

The covariance matrix of the residual for the estimate (25) is

$$Q = \mathbb{E} \left[\left(\hat{r} - \hat{R}\hat{\theta} \right) \left(\hat{r} - \hat{R}\hat{\theta} \right)^T \right].$$

Then, the estimate can be refined by suitably weighting the regressors, that is, by substituting (25) with

$$\hat{\theta}_D^Q = \left(\hat{R}^T \hat{Q}^{-1} \hat{R} \right)^{-1} \left(\hat{R}^T \hat{Q}^{-1} \hat{r} \right). \quad (27)$$

where \hat{Q} is a sample estimate of Q (see Stoica and Jansson (2000) for further details). Doing so, the covariance of the parameters becomes

$$\text{cov}(\hat{\theta}_D^Q) = (R^T Q^{-1} R)^{-1},$$

therefore a proper choice of Q might largely improve the statistical efficiency of the estimate.

6. SIMULATION EXAMPLE

To show the effectiveness of the proposed approach, the multivariable PI controller tuning problem for a LV100 gas turbine engine (see Yeddanapudi and Potvin (1997)) is considered as a simulation setup. The same example has been already employed in Miskovic et al. (2007), Hjalmarsson (1999) and Formentin et al. (2012) to illustrate the performance of other data-driven approaches, namely IFT, ICbT and MIMO VRFT.

The plant has 5 states: gas generator spool speed, power output, temperature, fuel flow actuator level and variable area turbine nozzle actuator level. The input signals are fuel flow and variable area turbine nozzle and the output signals are gas generator spool speed and temperature. The measurement noise is zero mean white noise with variance $0.0025I$. The reference model is selected as

$$M(q^{-1}) = \begin{bmatrix} \frac{0.4q^{-1}}{1 - 0.6q^{-1}} & 0 \\ 0 & \frac{0.4q^{-1}}{1 - 0.6q^{-1}} \end{bmatrix}.$$

In order to tune a multivariable coupled PI controller with the proposed non-iterative approach, an experiment has been set up by employing two PRBS input sequences u_1 and u_2 as described at the end of Section 4. First, the sequence u_1 is used to feed the first channel, then the same input is switched to the second channel. Finally, u_2 is used to separately feed the two channels, analogously to what has been done for u_1 . We note that these PRBS sequences satisfy the assumption in Theorem 1. The overall experiment can be considered as a unique dataset that can be used to design the MIMO controller with MIMO VRFT, so that the controllers obtained with the two methods can be (fairly) compared based on the same number of data. To select the length of the instrumental variable l , the order of magnitude of the (approximate) length of the impulse response of the element of M with the highest settling time is considered. Then, in this case, $l = 35$ if found by cross-validation. This approach is based on semi-empirical observations and has already been employed several times in the literature, see, *e.g.*, Formentin et al. (2013, 2014). The implementation of the discussed method returns the following transfer matrix

$$K_{dd}(q^{-1}) = \begin{bmatrix} \frac{0.3325 - 0.1154q^{-1}}{1 - q^{-1}} & \frac{0.4453 - 0.3532q^{-1}}{1 - q^{-1}} \\ \frac{15.33 - 14.88q^{-1}}{1 - q^{-1}} & \frac{-2.316 + 1.347q^{-1}}{1 - q^{-1}} \end{bmatrix}.$$

If the same dataset is used for MIMO VRFT design, the following controller is obtained instead:

$$K_{VRFT}(q^{-1}) = \begin{bmatrix} \frac{0.2079 + 0.06967q^{-1}}{1 - q^{-1}} & \frac{0.4494 - 0.3516q^{-1}}{1 - q^{-1}} \\ \frac{19.13 - 18.56q^{-1}}{1 - q^{-1}} & \frac{-3.145 + 2.241q^{-1}}{1 - q^{-1}} \end{bmatrix}.$$

Notice that the above K_{VRFT} is slightly different from the one given in Formentin et al. (2012). This is due

to the fact that the currently employed dataset uses different realizations of noise and inputs from the ones in Formentin et al. (2012). The reason is that the noise realizations and the input leading to the controller tuned in Formentin et al. (2012) cannot be used in the new approach. Therefore, to establish a fair comparison with the controller of Formentin et al. (2012) a different set of data, like the one employed in this section, needs to be used. Nonetheless, it should also be noticed that the quality of the step response of K_{VRFT} is the same of that obtained in Formentin et al. (2012) in terms of overshoot, performance in channel decoupling, settling time and overall shape.

A closed-loop noiseless experiment with the controller given by the proposed approach and the one returned by MIMO VRFT is illustrated in Figure 1. Notice that both the controllers yield good performance after being tuned with a single set of I/O data collected in open-loop operation.

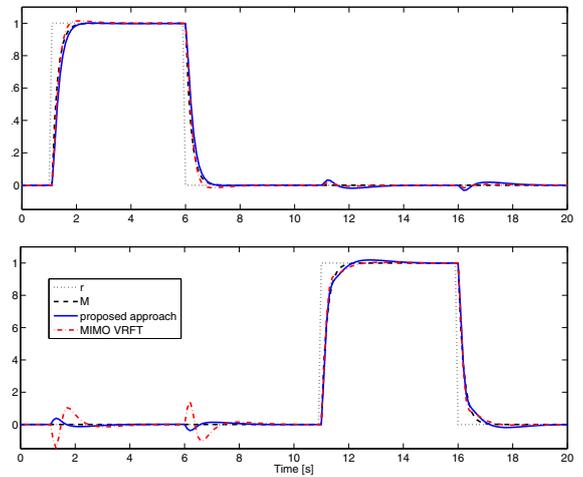


Fig. 1. Comparison of controllers tuned with the proposed method and with MIMO VRFT on noisy open-loop data and then applied in closed loop in noise-free context. Reference signals (dotted), ideal response (dashed) and achieved responses: MIMO VRFT controller (dash-dotted) and proposed approach (solid).

However, since MIMO VRFT returns only an approximated solution of Problem 2 (due to the fact that the optimal filter cannot be computed), the matching of M is less accurate than that given by the proposed approach, which is instead asymptotically optimal. In this example, this is especially visible in the coupling between the first input and the second output.

To give an overview of the data-driven solutions for multivariable control, consider now the performance of the above controllers together with the ones given by IFT and ICbT for the same example. In particular, Hjalmarsson (1999) provides, after 6 iterations and 30 experiments, the IFT controller

$$K_{IFT}(q^{-1}) = \begin{bmatrix} \frac{0.248 - 0.03q^{-1}}{1 - q^{-1}} & \frac{0.38 - 0.199q^{-1}}{1 - q^{-1}} \\ \frac{16.47 - 15.91q^{-1}}{1 - q^{-1}} & \frac{0.063 - 0.054q^{-1}}{1 - q^{-1}} \end{bmatrix}$$

whereas Miskovic et al. (2007) gives, after 8 iterations and 8 experiments, the ICbT controller

Table 1. Achieved cost J and MSE for different data-driven methods.

	proposed method	MIMO VRFT	ICbT	IFT
$J (\times 10^{-2})$	3.9	4.1	3.7	8.1
MSE	0.189	0.257	0.089	0.945

$$K_{ICbT}(q^{-1}) = \begin{bmatrix} \frac{0.3636 - 0.09866q^{-1}}{1 - q^{-1}} & \frac{0.3653 - 0.2691q^{-1}}{1 - q^{-1}} \\ \frac{18.69 - 18.16q^{-1}}{1 - q^{-1}} & \frac{-3.453 + 2.652q^{-1}}{1 - q^{-1}} \end{bmatrix}$$

In Table 1, the cost functions (5b) for the closed-loop system with each controller are reported.

To better appreciate the differences between the methods, the mean squared errors

$$MSE = \frac{1}{N_{cl}} \sum_{t=1}^{N_{cl}} \|y(t) - M(q^{-1})r(t)\|^2$$

are also computed over different noiseless step tests of the type in Figure 1 and reported in Table 1. In this expression of MSE , $r(t)$ denotes the step excitation, $\|\cdot\|$ denotes the Euclidean norm and N_{cl} is the number of samples of the closed-loop step test.

From the above results, it is clear that ICbT is the best method (IFT ended up with a local minimum), but the proposed approach allows one to reduce the MSE of 26% with respect to the only other non-iterative solution, *i.e.*, MIMO VRFT.

In terms of *hints for the user*, this means that, in case of costly experiments, the proposed solution seems to be the best choice, whereas, in case of cheap data collection, the proposed method can still be useful as an initial guess to start the iterative procedure of ICbT.

7. CONCLUSIONS

In this paper, a data-driven method for multivariable controller design has been presented. The method does not require the physical modeling of the process to control and is suited for fixed-parameterization coupled controllers. Simple least squares techniques based on instrumental variables are sufficient to obtain a consistent estimate of the optimal controller. Unlike the state of the art method for non-iterative data-driven MIMO control design, namely MIMO VRFT, the proposed method does not rely on a suboptimal filter, but it is asymptotically exact. Due to this fact, the accuracy of the closed-loop model matching can be enhanced, as illustrated on a benchmark simulation example.

Future work will be devoted to the optimal selection of the reference model and experimental validation of the proposed approach.

REFERENCES

K.J. Astrom and B. Wittenmark. *Adaptive control*. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, 1994. ISBN 0201558661.

A. Bunse-Gerstner, D. Kubalinska, G. Vossen, and D. Wilczek. H_2 -norm optimal model reduction for large scale discrete dynamical MIMO systems. *Journal of computational and applied mathematics*, 233(5):1202–1216, 2010. ISSN 0377-0427.

M.C. Campi, A. Lecchini, and S.M. Savaresi. Virtual reference feedback tuning: a direct method for the design

of feedback controllers. *Automatica*, 38(8):1337–1346, 2002.

S. Formentin and S.M. Savaresi. Noniterative data-driven design of multivariable controllers. In *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pages 5106–5111, 2011.

S. Formentin, S.M. Savaresi, and L. Del Re. Non-iterative direct data-driven controller tuning for multivariable systems: theory and application. *IET Control Theory & Applications*, 6(9):1250–1257, 2012.

S. Formentin, A. Karimi, and S.M. Savaresi. Optimal input design for direct data-driven tuning of model-reference controllers. *Automatica*, 49(6):1874–1882, 2013.

S. Formentin, K. Heusden, and A. Karimi. A comparison of model-based and data-driven controller tuning. *International Journal of Adaptive Control and Signal Processing*, 28(10):882–897, 2014.

M. Gevers. Identification for control: From the early achievements to the revival of experiment design. *European Journal of Control*, 11(4):335–352, 2005.

H. Hjalmarsson. Efficient tuning of linear multivariable controllers using iterative feedback tuning. *International journal of adaptive control and signal processing*, 13(7):553–572, 1999. ISSN 1099-1115.

H. Hjalmarsson. From experiment design to closed-loop control. *Automatica*, 41(3):393–438, 2005.

H. Hjalmarsson, M. Gevers, S. Gunnarsson, and O. Lequin. Iterative feedback tuning: theory and applications. *Control Systems Magazine, IEEE*, 18(4):26–41, 2002. ISSN 0272-1708.

J. Maciejowski. *Multivariable feedback design*. Addison-Wesley Wokingham, 1989.

L. Miškovic, A. Karimi, and D. Bonvin. Iterative controller tuning by minimization of a generalized decorrelation criterion. In *13th IFAC Symp. on System identification*, pages 1177–1182, 2003.

L. Miskovic, A. Karimi, D. Bonvin, and M. Gevers. Correlation-based tuning of decoupling multivariable controllers. *Automatica*, 43(9):1481–1494, 2007. ISSN 0005-1098.

M. Nakamoto. An application of the Virtual Reference Feedback Tuning method for multivariable process control. *Transactions of the Society of Instrument and Control Engineers*, 39(10):924–932, 2003.

T. Oomen, R. van der Maas, C.R. Rojas, and H. Hjalmarsson. Iterative data-driven H-infinity norm estimation of multivariable systems with application to robust active vibration isolation. *IEEE Transactions on Control Systems Technology*, 22(6), 2014.

S.M. Savaresi and G.O. Guardabassi. Approximate I/O feedback linearization of discrete-time non-linear systems via virtual input direct design. *Automatica*, 34(6):715–722, 1998.

S. Skogestad and I. Postlethwaite. *Multivariable feedback control: analysis and design*. Wiley New York, 2005. ISBN 0471942774.

P. Stoica and M. Jansson. MIMO system identification: state-space and subspace approximations versus transfer function and instrumental variables. *IEEE Transactions on Signal Processing*, 48(11):3087–3099, 2000.

M. Yeddanapudi and A.F. Potvin. *Nonlinear Control Design Blockset: User’s Guide*. Natick, MA: The Mathworks Inc, 1997.