Sparse Data-Driven Identification of Nonlinear Controllers ⋆

Kauã M. Antunez * **Alexandre S. Bazanella** *

* *Programa de Pós-Graduação em Engenharia Elétrica (PPGEE), UFRGS, (e-mails: kaua.antunez@ufrgs.br, bazanella@ufrgs.br).*

Abstract: This work presents the use of ℓ_1-regularization on the nonlinear formulation of the Virtual Reference Feedback Tuning. When the controller has a substantial quantity of parameters to be estimated, which tends to be the case in black-box nonlinear identification, the least-squares method yields estimates with inadequate statistical properties. To handle that, the use of ℓ_1-regularization on the controller estimation is proposed, reducing the variance and the bias, as well as thresholding the unneeded controller parameters. In this paper three different regularization methods are described and their algorithms are presented. For the purpose of illustrating the main properties of these methods, two numerical examples are presented.

Keywords: Data-driven Control; Regularization; System Identification; Nonlinear Systems.

1. INTRODUCTION

The control of dynamical systems is usually made through the identification of the process, followed by the controller's design based on the model thus obtained and on the performance requirements. In the Data-Driven (DD) framework the ambition is to estimate a controller that approximates the dynamical behavior of the process to a specified Reference Model, without the intermediate step of identifying a model (Bazanella et al., 2012).

In the midst of the DD methods, one of the most successful is the *Virtual Reference Feedback Tuning* (VRFT) Campi et al. (2002). This methodology is part of the direct methods group, i.e. a single batch of input-output data is requisite to tune the controller. Several extensions for the VRFT have been researched and presented in the literature: Lecchini et al. (2002) with a 2-degree of freedom approach, Campestrini et al. (2011) to deal with Non-minimum Phase (NMP) zeros of the plant, Campestrini et al. (2016) which shows the MIMO case, and Campi and Savaresi (2006) for the nonlinear scenario. In addition, there is the iterative methods group, where the most known and the pioneer is the Iterative Feedback Tuning (IFT) Hjalmarsson et al. (1998). The must of iterative methods is the sequence of experiments to improve the controller's parameters.

As a result of VRFT being a one-shot design and the controller possessing a fixed configuration, the optimization problem can be solved via the Least Squares (LS) or the Instrumental Variables (IV) method. However, the fact that theses methods display poor statistical properties is a well-known fact, so that alternative approaches for the optimization are still being sought Garcia and Bazanella (2020).

⋆ This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001 and by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Consequently, as the estimates exhibits poor statistical proprieties, the system's closed-loop performance is directly affected. Hence, the method becomes less attractive to be employed in real industrial applications where high noise levels are present. Also, it is a well-known fact that the majority of the systems show a nonlinearity, though this aspect is often ignored in the control design.

Inside the DD community, there are fundamental studies about the design of nonlinear controllers aiming at a closed-loop system with a linear behavior. These works usually involve a considerable knowledge about the process and the applied signal for achieving satisfactory estimates Campi and Savaresi (2006). In Bazanella and Neuhaus (2014), a new class of controllers is estimated through the VRFT method: the rational and polynomial structures. This rational structure is able to represent several real systems, as it uses the previous input and output signals, however the algorithm used to estimate a rational nonlinear system is a sequence of Least-Squares. Besides, the work in Bazanella and Neuhaus (2014) produces a controller with a complex structure, since it is does not apply any ℓ_1-regularization technique.

In the DD framework, often there is very little - or no - prior information on the process available. Under these circumstances, an overparameterized controller structure is required. Therefore, a new design tool is essential to guarantee the best statistical proprieties possible. In the linear monovariable - Single-Input Single-Output (SISO) - context, the addition of the ℓ_2-regularization has been discussed in Rallo et al. (2016); Formentin and Karimi (2014), where authors include the regularization to reduce the estimates covariance using instrumental variables and enhance the system's performance. Likewise, the work Boeira (2018) presents a *Bayesian* perspective for the multivariable VRFT method.

With these ideas in mind, this work's main idea is to exploit the ℓ_1-regularization to get better estimates on overparameterized controllers and yielding sparse ones. The contributions of this paper are: comparing the already proposed ℓ_1-regularization methods in the literature with the one developed herein, analyze the noise effect and the input signal on the estimates' quality.

2. NONLINEAR VRFT

Given the following open-loop Single-Input Single-Output (SISO) system

$$y(t) = \mathcal{P}u(t) + H(q)\epsilon(t), \tag{1}$$

where $\mathcal{P}$ is the nonlinear process, $H(q)$ is the noise model, $y(t)$ and $u(t)$ are the process' input and output signals, respectively, $\epsilon(t)$ is a zero-mean white noise signal with variance σ^2, and q is the forward-shift operator, i.e. $qu(t) = u(t+1)$. The controller is linearly parameterized using a linear class of controllers and a library of nonlinear functions.

$$\begin{aligned} u(t) &= \mathcal{C}(v(t), \rho) \\ &= \rho^T \Phi(v(t)), \end{aligned} \tag{2}$$

where $v(t) = \bar{C}(q)e(t)$, with $\bar{C}(q)$ being the linear part of the controller, ρ is the parameters vector, and $\Phi(v(t))$ is the regressor matrix formed by the library of nonlinear function ϕ.

The Model Reference based design aims to solve the following optimization problem

$$\rho^* = \arg\min_{\rho} J_y(\rho), \tag{3}$$

$$J_y(\rho) \triangleq \bar{E}[y(t,\rho) - y_d(t)]^2, \tag{4}$$

where $\bar{E}[x(t)] = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} E[x(t)]$, and $E[.]$ is the expected value. $J_y(\rho)$ is the reference tracking performance criterion, $y(t,\rho)$ is the closed-loop system output signal obtained with the controller $\mathcal{C}(v(t),\rho)$, whereas $y_d(t)$ is the output collected through the ideal controller $\mathcal{C}_d(v(t))$. Particularly, $y_d(t) = T_d(q)r(t)$, $T_d(q)$ is the reference model and $r(t)$ is the reference signal. Hence, the ideal controller is the one that achieves $T(q,\rho) = T_d(q)$. In other words, $\overline{r}(t)$ is the input applied to the reference model that would generate the measured output $y(t)$.

Assumption 2.1. Matching condition - $C_d \in \mathbb{C}$

$$\exists \rho^* : \mathcal{C}(v(t),\rho^*) \equiv \mathcal{C}_d(v(t),\rho) \tag{5}$$

This condition is reciprocal with the assumption that the system is in the model set.

The complexity of solving (4) is that it depends on the unknown plant and it is a non-convex function. The VRFT approach remodels this optimization problem by using the Virtual Reference (VR). Firstly, a sufficiently rich signal $u(t)$ is applied to the plant and the output $y(t)$ is measured, then the *virtual reference* $\overline{r}(t)$ is determined as a result of $\overline{r}(t) = T_d^{-1}(q)y(t)$. Figure 1 illustrates the experiment and the mentioned signals. In this work, we propose an extension of the linear signal, designated as $v(t)$, through a library of nonlinear functions defined as $\phi(v(t))$ as in (6).

$$\phi(v(t)) = \begin{bmatrix} | & | & | & | & | \\ \phi_1(v(t)) & \phi_2(v(t)) & \phi_3(v(t)) & \ldots & \phi_n(v(t-N)) \\ | & | & | & | & | \end{bmatrix} \tag{6}$$

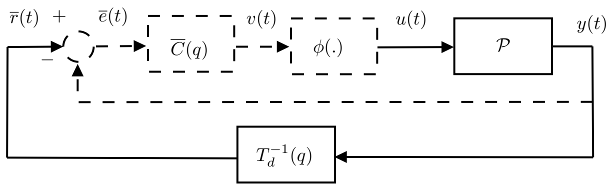

Figure 1. Virtual reference feedback experiment.

Therefore, the VRFT goal is to solve the following optimization problem

$$\rho^* = \arg\min_{\rho} J^{VR}(\rho), \tag{7}$$

$$\begin{aligned} J^{VR}(\rho) &\triangleq \bar{E}[u(t) - \mathcal{C}(v(t),\rho)]^2 \\ &= \bar{E}[u(t) - \rho^T\Phi(v(t))]^2. \end{aligned} \tag{8}$$

The objective function shown in (8) has the same global minimum as in (4), this is the foundation of the VRFT. However, it can be solved through the known least-squares (LS) method, as consequence of structure a controller which is linear in parameters. The advantages of the VRFT method are that it can be solved only using the measured data; it is a direct method (only one batch of data are needed).

In this paper we propose the use of a library of nonlinear functions generated as a truncated Taylor series of the ideal nonlinear function. Given that the process is unknown, with little or no prior information about it, we do not know a priori the appropriate order of the truncation or which terms are necessary to obtain a good approximation. A high order controller with all terms of the Taylor series present must then be assumed and some technique can be applied to reduce the controller complexity, yielding a *sparse nonlinear controller.*

3. VRFT WITH ℓ_1-REGULARIZATION

In this subsection is depicted the ℓ_1-regularization methods applied to identify the sparse controllers. The large variance effect on the least-squares identification due to large number of parameters is a well-known problem in the system identification framework. In such situations the estimated parameters may lead to a poor closed-loop performance or even to an unstable closed-loop.

In the linear SISO and Multiple-Input Multiple-Output MIMO VRFT backgrounds the ℓ_2-regularization has been proposed to lower the variance compared to the Instrumental Variables (IV) method (Boeira and Eckhard, 2019) (Formentin and Karimi, 2014) (Rallo et al., 2016), even though the controller's structure was not sparse. Herein, the interest is to not only reduce the Mean Squared Error

(MSE) per se, but also the number of identified parameters - which will also imply a reduction of the variation for the remaining parameters. In order to obtain this sparse controller, it is recommended to include the ℓ_1-regularization term and the ℓ_1 penalty term in the objective function (8) Brunton et al. (2016). To delimit the ℓ_1-norm of the parameters prompts a lower model complexity, therefore decreasing the variance.

$$\rho^*_{Reg} = \arg\min_{\rho_{Reg}} J^{VR}_{Reg}(\rho), \tag{9}$$

$$J^{VR}_{Reg}(\rho) \triangleq J^{VR}(\rho) + \lambda \sum_{j=1}^{P} |\rho_j| \tag{10}$$

with λ being a scalar and entitled the *regularization parameter*, and $\rho \in \mathbb{R}^n$. The regularization parameter weighs the ℓ_1 penalty, i.e. as $\lambda \to \infty$ $\rho^* \to 0$.

The point in question is how to solve (10) with the purpose of reducing the MSE and the quantity of parameters. In this work we address three different approaches to solve (10): the well-known Least Absolute Shrinkage and Selection Operator (LASSO) (Tibshirani, 1996), the Sequential Thresholded Least-Squares (STLS) (Brunton et al., 2016), and the Sequential Thresholded Least-Squares 2 (STLS_2), the last being introduced in this work.

3.1 LASSO

LASSO is a famous method for regularizing the least-squares using the ℓ_1-regularization achieving a sparse solution. The algorithm to solve this problem (11) is similar to the Alternating Direction Method of Multipliers (ADMM) (Boyd et al., 2011).

$$\begin{array}{l} min \ \left\|u(t) - \rho^T\Phi(t))\right\|_2^2 + \lambda\|\sigma\|_1 \\ s.t. \ \ \rho - \sigma = 0 \end{array} \tag{11}$$

The ADMM algorithm is presented in (12),

$$\begin{array}{c} \rho_{k+1} = (\Phi^T\Phi + \xi I)^{-1}(\Phi u + \xi(\sigma_k - u_k)) \\ \sigma_{k+1} = S_{\lambda/\xi}(\rho_{k+1} + u_k) \\ u_{k+1} = u_k + \rho_{k+1} - \sigma_{k+1}, \end{array} \tag{12}$$

where Φ is the regressed data, ξ is the augmented Lagrangian parameter and $S(.)$ is the soft thresholding operator (13).

$$S_\lambda(z) = \begin{cases} z - \lambda & z > \lambda \\ 0 & |z| < \lambda \\ z + \lambda & z < \lambda \end{cases} \tag{13}$$

The LASSO disadvantage is that it requires a considerably computational effort to accomplish the solution. On the other hand, employing the *k-fold Cross Validation* (CV) provides a minimum variance and also a sparse estimate (James et al., 2013).

3.2 Sequential Thresholded Least Squares

Concerning the computational endeavor and with the same intention to find a sparse solution on overdetermined systems, the Sequential Thresholded Least Squares is proposed in (Brunton et al., 2016) to identify model parameters of nonlinear dynamical systems. In this work, we exploit the STLS to identity the controller's parameters.

The STLS algorithm applied herein is presented below:

Algorithm 1 Sequential Thresholded Least-Squares

Data: Reference Model $T_d(q)$, controller structure $\bar{C}(q)$, library ϕ, threshold λ_{STLS}, measured data ($u(t)$ and $y(t)$), $t = 1, ..., N$
Result: Estimated parameters ρ_{STLS}
Generate the virtual reference and the regressor matrix
$\overline{r}(t) = T_d^{-1}(q)y(t)$
$\overline{e}(t) = \overline{r}(t) - y(t)$
$v(t) = \overline{C}(q)\overline{e}(t)$
Generate the regressor matrix using the library ϕ
$\Phi = [\phi_1(v(t)) \ \phi_2(v(t)) \ ... \ \phi_n(v(t-N))]$
Search for the small parameters
Initial guess: least-squares $\rho = (\Phi^T\Phi)^{-1}\Phi^T u(t)$
Determine the ρ indexes less than λ_{STLS}
$\alpha = |\rho| \leq \lambda_{STLS}$
Threshold the parameters
$\rho_\alpha = 0$
Determine the ρ indexes greater than λ_{STLS}
$\beta = |\rho| > \lambda_{STLS}$
Regress the dynamics onto remaining terms
$\rho_{STLS} = (\Phi_\beta^T\Phi_\beta)^{-1}\Phi_\beta u(t)$

where $\alpha \in \mathbb{R}^p$, with p being the number of zero parameters and $\beta \in \mathbb{R}^q$, with q being the number of nonzero parameters.

3.3 Sequential Thresholded Least Squares 2

Finding an appropriate threshold in the STLS just described is a critical task for which there seems to be no firm guidelines in the literature. It is doubtful whether such firm guidelines can ever be derived, since a single threshold must be applied to parameters with hugely different units. Thus, it seems seem wiser to evaluate the parameters whose net contribution to the objective function (8) is smaller than a threshold instead; specifically, $\sum_{j=1}^{M} |\rho_j\phi_j(t)| < \lambda$. This is what we propose here, under the name Sequential Thresholded Least Squares 2 (STLS_2).

The (STLS_2) algorithm is defined as follows:

4. NUMERICAL EXAMPLES

This section's goal is to illustrate the efficiency of the ℓ_1 regularization methods compared do the least-squares.

4.1 Hammerstein System

The first case study is implemented with a Hammerstein System, where the linear part of the open-loop process is given by

$$G(q) = \frac{0.2}{q - 0.8},$$

and the static nonlinearity is a $\sqrt{|.|}$ and $H(q) = 1$

ISSN: 2175-8905

DOI: 10.20906/sbai.v1i1.2592

Algorithm 2 Sequential Thresholded Least-Squares 2

Data: Reference Model $T_d(q)$, controller structure $\bar{C}(q)$, library ϕ, threshold λ_{STLS_2}, measured data ($u(t)$ and $y(t)$), $t = 1, ..., N$
Result: Estimated parameters ρ_{STLS_2}
Generate the virtual reference and the regressor matrix
$\overline{r}(t) = T_d^{-1}(q)y(t)$
$\overline{e}(t) = \overline{r}(t) - y(t)$
$v(t) = \overline{C}(q)\overline{e}(t)$
Generate the regressor matrix using the library ϕ
$\Phi = [\phi_1(v(t))\ \phi_2(v(t))\ ...\ \phi_n(v(t-N))]$
Search for the small parameters
Initial guess: least-squares $\rho = (\Phi^T\Phi)^{-1}\Phi^T u(t)$
Determine the ρ indexes contribution less than λ_{STLS_2}
$\alpha = |\rho_j\phi_j| \le \lambda_{STLS_2}$
Threshold the parameters
$\rho_\alpha = 0$
Determine the ρ indexes greater than λ_{STLS_2}
$\beta = |\rho| > \lambda_{STLS_2}$
Regress the dynamics onto remaining terms
$\rho_{STLS} = (\Phi_\beta^T\Phi_\beta)^{-1}\Phi_\beta u(t)$

The desired closed-loop performance chosen for the system is given by the following transfer function

$$T_d(q) = \frac{0.3}{q - 0.7}.$$

In the linear case, the ideal controller $C_d(q)$ would be the following

$$C_d(q) = [1.5\ \ 0.3]\left[1\ \ \tfrac{1}{q-1}\right]^T,$$

which is a Proportional-Integral (PI) controller. The controller class $\mathbb{C}$ chosen is PI controller as well. In such manner, the matching condition is met.

In the nonlinear case, it is clear that the ideal controller would be PI in addition to the inverse of the nonlinearity that is in the process, i.e. $f(.) = (.)^2$. The expansion of the linear signals $v_p(t)$ and $v_i(t)$ was made up to the third order, thus generating 15 regressors vectors:

$$\Phi = \left[v_p(t)\ v_p^2(t)\ v_p^3(t)\ v_i(t)\ v_i^2(t)\ v_i^3(t)\ v_p(t)v_i(t)\ ...\ v_p^3(t)v_i^3(t)\right].$$

Thus the ideal controller would have the following parameters

$$\begin{aligned}\rho_0^T &= \left[0\ K_p^2\ 0\ 0\ K_i^2\ 0\ 2K_pK_i\ \ \ 0\right]^T \\ &= \left[0\ 2.25\ 0\ 0\ 0.09\ 0\ 0.9\ ...\ 0\right]^T,\end{aligned}$$

so, the ideal controller would be

$$\mathcal{C}_d(v(t)) = \mathcal{C}(v(t), \rho) = \rho_0^T\Phi(v(t)),$$

with $v_p(t) = e(t)$ and $v_i(t) = \dfrac{1}{q-1}e(t)$.

The input signal $u(t)$ employed to excite the plant was a Pseudo Random Binary Signal (PRBS) multiplied by the absolute value of a Gaussian noise with zero mean and variance $\sigma^2 = 1$, with $N = 1500$ samples. Besides, the plant's output is affected by a gaussian noise with variance $\sigma_e^2 = 1\times10^{-4}$. Concerning the LASSO algorithm, the MATLAB function *lasso* was used. The regularization parameter λ_{LASSO} was calculated through the 10-fold Cross Validation algorithm so that it would yield minimum variance, and $\lambda_{STLS} = 0.05$ and $\lambda_{STLS_2} = 20$ (which corresponds approximately to 1% of the contribution to the objective function).

To evaluate the proposed technique, 100 Monte Carlo simulations were run with distinct noise realizations. The major objective of inserting the regularization on the VRFT was to draw a better closed-loop performance. This evaluation was done through the objective function $J_y(\rho)$, in addition to the sum of all the estimated zeros in each Monte Carlo simulation.

Table 1 exhibits the average controller gains that should be the nonzero parameters.

Table 1. Average Estimated Parameters

Regressor	$\hat{E}(\rho_{LS})$	$\hat{E}(\rho_{LASSO})$	$\hat{E}(\rho_{STLS})$	$\hat{E}(\rho_{STLS_2})$	ρ^*
v_p^2	1.9895	2.1496	2.0576	2.2177	2.25
v_p^2	0.0818	0.0869	0.0858	0.0893	0.09
$v_p.v_i$	0.9595	0.8604	0.8769	0.9000	0.9

To evaluate the sparsity of the estimation along all the Monte Carlo Simulations, we present Table 2, which contains the number of estimated zeros by the four methods and the ideal quantity as well.

Table 2. Total number of zeros

Method	N_0
LS	0
LASSO	661
STLS	640
$STLS_2$	625
Ideal	1200

If now we turn to the interpretation of the objective function, through Table 3 it is evident that all the regularization methods surmount the classical VRFT with the Least-Squares. The cost $\hat{J}_y(\hat{E}(\rho_{LS}))$ is 25% worst than the minimum, while the LASSO achieves the minimum up to three correct significant digits.

The estimate of this function is obtained through

$$\hat{J}_y(\hat{E}(\rho)) = \frac{1}{N}\sum_{t=1}^{N}(y(t,\rho) - y_d(t))^2.$$

Table 3. Objective Function Estimation

$\hat{J}_y(\hat{E}(\rho_{LS}))$	0.1586×10^{-3}
$\hat{J}_y(\hat{E}(\rho_{LASSO}))$	0.1268×10^{-3}
$\hat{J}_y(\hat{E}(\rho_{STLS}))$	0.1325×10^{-3}
$\hat{J}_y(\hat{E}(\rho_{STLS_2}))$	0.1286×10^{-3}
$\hat{J}_y(\rho^*)$	0.1265×10^{-3}

Analyzing the boxplots in Figure 2, it is possible to confirm that the ℓ_1-regularization methods decreased both the variance and the bias of the estimate, with the LASSO presenting the best results. If we draw the attention to the STLS and $STLS_2$ methods, they attained a worst variance compared to the LASSO.

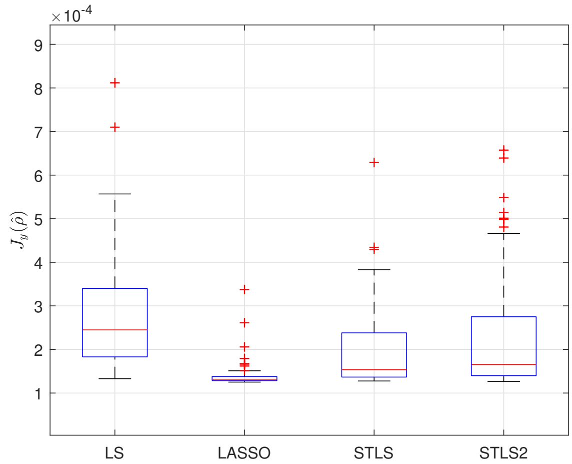

Figure 2. Comparison of $J_y(\hat{\rho})$ for the classical and Regularized VRFT.

4.2 Continuous Stirred-Tank Reactor

This subsection presents another case study: the *Continuous Stirred-Tank Reactor* (CSTR). (14) is a representation of the CSTR.

$$\mathcal{P}: \begin{cases} \dot{x_1}(t) = -2x_1^2(t) + (1 - x_1(t))u(t) \\ \dot{x_2}(t) = x_1^2(t) - x_2(t)u(t), \\ y(t) = x_2(t) \end{cases} \quad (14)$$

as in the last example, the noise model is $H(q) = 1$.

The desired closed-loop performance chosen for the system is given by the transfer function

$$T_d(q) = \frac{0.0216(q + 0.8)}{(q - 0.85)^2}.$$

The selected controller class is a Proportional-Integral-Derivative (PID)

$$\bar{C}(q) = \begin{bmatrix} 1 & \frac{1}{q-1} & \frac{q-1}{q} \end{bmatrix},$$

it is uncomplicated to conclude that $C_d(q) \notin \mathcal{C}$.

The expansion of the linear signals $v_p(t)$, $v_i(t)$ and $v_d(t) = e(t)\dfrac{q-1}{q}$ was made up to the second order, generating 23 parameters to be estimated.

The input signal $u(t)$ applied into the plant is a sequence of steps to yield an output $y(t)$ in the range from 0.1 to 0.4. Furthermore, the plant's output is affected by a Gaussian noise with variance $\sigma^2 = 2.5 \times 10^{-7}$. As with the first example, we ran 100 Monte Carlo Simulations. The regularization parameter λ_{LASSO} is calculated as previously, the λ_{STLS} is the average of ρ_{LS} for each noise realization and λ_{STLS_2} is the average of contribution to the objective function for each noise realization as well.

Examining the Table 4, we can observe that the LASSO method presented the best sparse identification, i.e. the majority number of total zeros. After, the proposed method $STLS_2$ presents 1499 identified zeros which is near to the LASSO.

Table 4. Total number of zeros

Method	N_0
LS	0
LASSO	1531
STLS	944
$STLS_2$	1499

If now we turn to the interpretation of the objective function, through Table 5 it is evident that all the regularization methods surmount the classical VRFT with the Least-Squares.

Table 5. Objective Function Estimation

$\hat{J}_y(\hat{E}(\rho_{LS}))$	0.1625×10^{-4}
$\hat{J}_y(\hat{E}(\rho_{LASSO}))$	0.0905×10^{-4}
$\hat{J}_y(\hat{E}(\rho_{STLS}))$	0.0305×10^{-4}
$\hat{J}_y(\hat{E}(\rho_{STLS_2}))$	0.0422×10^{-4}

Analyzing the boxplots in Figure 3, it is possible to confirm that the ℓ_1-regularization methods decreased the variance, with the LASSO presenting the best results. If we draw the attention to the STLS and $STLS_2$ methods, they attained an excellent minimum, with the $STLS_2$ overcoming the STLS in terms of variance.

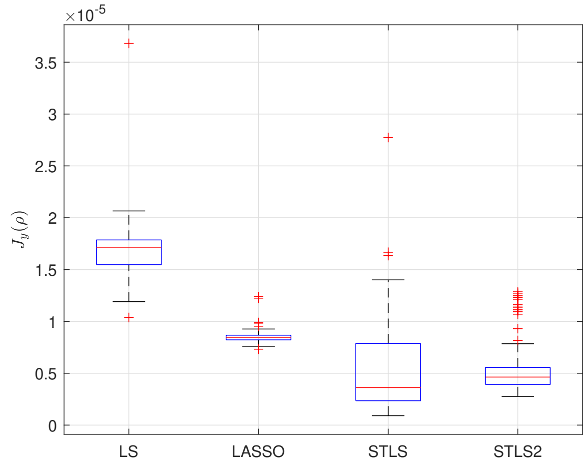

Figure 3. Comparison of $J_y(\hat{\rho})$ for the classical and Regularized VRFT.

To illustrate the closed-loop performance in the designated range, we show Figure 4. It can be seen that the regularization methods performed better that the Least-Squares.

5. CONCLUSIONS

This paper present the implementation of the VRFT to estimate sparse nonlinear controllers. The drawback of using the classical VRFT with the Least Squares method is that it is not able to find sparse solutions. Thus, we presented three different ℓ_1-regularization methods, originally used in system identification, to estimate this solutions using the data-driven controller design problem. The regularization methods benefits were illustrated through two different study cases, which showed that both the variance and the bias have improved significantly, resulting in much better closed-loop performance in most cases. Future works will be focused on how to find an even more

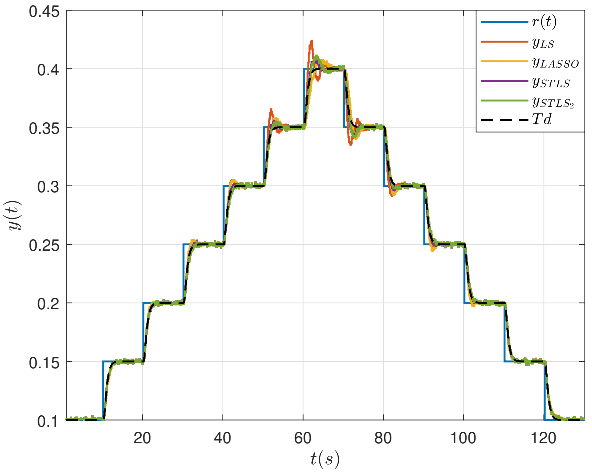

Figure 4. CSTR closed-loop performance.

sparse estimate and the application of the cross-validation techniques in the STLS and $STLS_2$ methods.

REFERENCES

Bazanella, A.S., Campestrini, L., and Eckhard, D. (2012). *Data-driven controller design: the H2 approach*. Springer Science & Business Media.

Bazanella, A.S. and Neuhaus, T. (2014). Tuning nonlinear controllers with the virtual reference approach. *IFAC Proceedings Volumes*, 47(3), 10269–10274.

Boeira, E.C. (2018). *Sintonia De Controladores Multivariáveis Pelo Método Da Referência Virtual Com Regularização Bayesiana*. Master thesis, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brasil.

Boeira, E.C. and Eckhard, D. (2019). Multivariable virtual reference feedback tuning with bayesian regularization. *Anais da Sociedade Brasileira de Automática*, 1(1).

Boyd, S., Parikh, N., and Chu, E. (2011). *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc.

Brunton, S.L., Proctor, J.L., and Kutz, J.N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences*, 113(15), 3932–3937.

Campestrini, L., Eckhard, D., Chía, L.A., and Boeira, E. (2016). Unbiased MIMO VRFT with application to process control. *Journal of Process Control*, 39, 35–49.

Campestrini, L., Eckhard, D., Gevers, M., and Bazanella, A.S. (2011). Virtual Reference Feedback Tuning for Non-Minimum Phase Plants. *Automatica*, 47(8), 1778–1784. doi:10.1016/j.automatica.2011.04.002.

Campi, M.C., Lecchini, A., and Savaresi, S.M. (2002). Virtual Reference Feedback Tuning: A Direct Method for the Design of Feedback Controllers. *Automatica*, 38(8), 1337–1346. doi:10.1016/S0005-1098(02)00032-8.

Campi, M.C. and Savaresi, S.M. (2006). Direct Nonlinear Control Design: The Virtual Reference Feedback Tuning (VRFT) Approach. *IEEE Transactions on Automatic Control*, 51(1), 14–27. doi:10.1109/TAC.2005.861689.

Formentin, S. and Karimi, A. (2014). Enhancing statistical performance of data-driven controller tuning via l2-regularization. *Automatica*, 50(5), 1514–1520.

Garcia, C.S. and Bazanella, A.S. (2020). Selection of informative intervals in routine operating data for use in data-driven control design. In *2020 IEEE Conference on Control Technology and Applications (CCTA)*, 225–230. IEEE.

Hjalmarsson, H., Gevers, M., Gunnarsson, S., and Lequin, O. (1998). Iterative Feedback Tuning: Theory and Applications. *IEEE Control Systems*, 26–41.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, volume 112. Springer.

Lecchini, A., Campi, M.C., and Savaresi, S.M. (2002). Virtual Reference Feedback Tuning for Two Degree of Freedom Controllers. *International Journal of Adaptive Control and Signal Processing*, 16(5), 355–371. doi:10.1002/acs.711.

Rallo, G., Formentin, S., Chiuso, A., and Savaresi, S.M. (2016). Virtual reference feedback tuning with bayesian regularization. In *2016 European control conference (ECC)*, 507–512. IEEE.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288.