DOI: 10.1109/ECC.2014.6862222

Design and Testing of a Constrained Data-Driven Iterative Reference Input Tuning Algorithm*

Mircea-Bogdan Radac, *Member, IEEE*, Radu-Emil Precup, *Senior Member, IEEE*, and Emil M. Petriu, *Fellow, IEEE*

Abstract—**This paper presents aspects concerning the design and testing of a new data-driven Iterative Reference Input Tuning (IRIT) algorithm that solves a reference trajectory tracking problem expressed as an optimization problem with control signal saturation constraints and control signal rate constraints. The design of the IRIT algorithm uses an experiment-based stochastic search algorithm formulated in the framework of Iterative Learning Control (ILC) in order to combine the advantages of data-driven control and of ILC. The iterative tuning is model-free in the sense it does not use control system models. A set of simulation results tests and validates the IRIT algorithm in a case study related to a representative mechatronics application that deals with the position control of a nonlinear aero-dynamical system. The IRIT algorithm offers the performance improvement by few iterations and experiments conducted on the process.**

I. Introduction

Data-driven optimization techniques for controller design offer the control system (CS) performance improvement using no a priori model information on the process or little such information using simple specifications in terms of easily interpretable performance indices. The performance indices are usually specified in the time domain (for example, the rise time, the overshoot, the settling time), and they are aggregated in general integral-type or sum-type objective functions (OFs) including the Linear Quadratic Gaussian (LQG) ones. The minimization of these functions in the framework of constrained optimization problems can fulfill different objectives such as reference trajectory tracking (including model reference tracking), control signal penalty, disturbance rejection, etc.

The main data-driven techniques that carry out the iterative experiment-based update of controller parameters are Iterative Feedback Tuning (IFT) [1], Simultaneous Perturbation Stochastic Approximation [2] and Model-free Control [3]. The most popular non-iterative technique is Virtual Reference Feedback Tuning (VRFT) [4]. These techniques use various approaches to ensure model-free controller tuning. However, the tuning to achieve reference trajectory tracking does not guarantee robust stability or robust performance.

*Manuscript received October 4, 2013. This work was supported by a grant of the Romanian National Authority for Scientific Research, CNCS – UEFISCDI, project number PN-II-ID-PCE-2011-3-0109, and by a grant from the NSERC of Canada.

M.-B. Radac and R.-E. Precup (corresponding author) are with the Department of Automation and Applied Informatics, Politehnica University of Timisoara, 300223 Timisoara, Romania (phone: +40 2564032 -29, -30, -26; fax: +40 2564032 -14; e-mail: mircea.radac@upt.ro; radu.precup@upt.ro).

E. M. Petriu is with the School of Electrical Engineering and Computer Science, University of Ottawa, 800 King Edward, Ottawa, ON, K1N 6N5 Canada (petriu@uottawa.ca).

The reference trajectory tracking can be considered as a reference input design over an initial CS with a priori tuned controllers in order to solve stability and disturbance rejection issues [5]. With this regard the iterative controller parameter update laws are replaced by reference input sequence update laws and the Iterative Learning Control (ILC) framework [6]–[8] is applied.

An experiment-based approach to the reference trajectory is given in [5] accounting for the control signal saturation constraints and employing an Interior Point Barrier (IPB) algorithm. This paper extends these results, and the main contributions with respect to the state-of-the-art are: a new data-driven Iterative Reference Input Tuning (IRIT) algorithm is offered, the operating conditions in the informative gradient experiments are restricted near nominal trajectories, the derivations for the inclusion of the constraints on the control signal rate are presented with the theory developed for Linear Time-Invariant (LTI) systems, we show that only one gradient experiment is needed in order to estimate the gradient of the penalty function that does not depend on the number of the constraints or on their type, a convincing mechatronics case study concerning a position CS for a nonlinear aero-dynamical system is treated, and the simulation results show that despite the LTI framework was used, the proposed ILC approach works well for the investigated system.

The advantages of these new contributions with respect to the previously analyzed literature are: the IRIT algorithm uses experiments conducted on the real-world CS. Therefore, it can compensate for process nonlinearities and uncertainties along the iterations and experiments, and it is model-free in the sense it does not use models of the controlled process, the IRIT algorithm employs a reduced number of experiments that enables cost-effective tuning, the handling of two categories of constraints in the optimization problems specific to reference trajectory tracking makes the IRIT-based CSs benefit of the advantages of predictive control, our approach is a special case of supervised learning that makes it belong to a more general context, with the relations of IRIT with supervised learning and with unsupervised learning discussed in [5], and due to its iterative nature, our algorithm can be started or stopped whenever necessary which is different from the adaptive control paradigm, the computations are carried out offline, so the computational effort is low.

DOI: 10.1109/ECC.2014.6862222

The paper is organized as follows. The next section presents the problem setting concerning the reference trajectory tracking problem solved by data-driven iterative optimization. Section III deals with the model-free estimation of OF's gradient with guaranteed convergence. Section IV proposes the model-free constrained optimal control problem and gives the formulation of the IRIT algorithm. Section V targets the validation of the algorithm. The conclusions are highlighted in Section VI. The computation of the gradients of the quadratic penalty functions with respect to the reference input vector is derived in the appendix.

II. Data-Driven Approach to Reference Trajectory Tracking

The discrete time LTI Single Input-Single Output (SISO) CS is characterized by

$$y(\boldsymbol{\rho}, r, k) = T(\boldsymbol{\rho}, q^{-1})r(k) + S(\boldsymbol{\rho}, q^{-1})v(k)\,, \tag{1}$$

where k is the discrete time argument, $y(k)$ is the process output sequence, $r(k)$ is the reference input sequence, $v(k)$ is the zero-mean stationary and bounded stochastic disturbance input sequence acting on the process output and accounting for various types of load or measurement disturbances. The parameter vector $\boldsymbol{\rho}$ contains the parameters of the controller. $S(\boldsymbol{\rho}, q^{-1})$ is the sensitivity function, $T(\boldsymbol{\rho}, q^{-1})$ is the complementary sensitivity function

$$\begin{aligned} &S(\boldsymbol{\rho}, q^{-1}) = 1/[1 + P(q^{-1})C(\boldsymbol{\rho}, q^{-1})], \\ &T(\boldsymbol{\rho}, q^{-1}) = 1 - S(\boldsymbol{\rho}, q^{-1}), \end{aligned} \tag{2}$$

$P(q^{-1})$ is the process transfer function, $C(\boldsymbol{\rho}, q^{-1})$ is the controller transfer function which is parameterized by the parameter vector $\boldsymbol{\rho}$ and q^{-1} is the one step delay operator.

In a reference trajectory tracking problem the sequence to be tracked by the output, i.e. the desired output $y^d(k)$, can be generated by a reference model. The control signal sequence $u(k)$ is not explicit in (1), but it may represent interest in control when constraints are required and the control effort should be manipulated.

An ILC framework [6]–[8] is applied next to the optimal reference input tuning for achieving trajectory tracking. For a relative degree n of the closed-loop CS transfer function $T(q^{-1})$, the lifted form representation for an N samples experiment length in the deterministic case is

$$\mathbf{Y} = \mathbf{T}\,\mathbf{R} + \mathbf{Y}_0\,, \tag{3}$$

with the matrices

$$\begin{aligned} &\mathbf{Y} = [y(n)\ \ldots y(N-1)]^T, \mathbf{R} = [r(0) \ldots r(N-n-1)]^T \\ &\mathbf{Y}_0 = [y_{10} \quad y_{20} \quad \ldots \quad y_{(N-n)0}]^T, \mathbf{T} = \begin{bmatrix} t_1 & 0 & \ldots & 0 \\ t_2 & t_1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ t_{N-n} & t_{N-n-1} & \ldots & t_1 \end{bmatrix}, \end{aligned} \tag{4}$$

$\mathbf{R}$ is the reference input vector which contains the reference input sequence over the time interval $0 \le k \le N-n-1$, $\mathbf{Y}$ is the controlled output vector, t_i is the i^{th} impulse response coefficient of $T(q^{-1})$, $\mathbf{T}$ is a lower-triangular Toeplitz matrix, $\mathbf{Y}_0$ is the free response of the CS due to nonzero initial conditions, and the superscript T indicates matrix transposition. Zero initial conditions can be assumed without loss of generality, and the tracking error vector $\mathbf{E}$ is

$$\mathbf{E} = \mathbf{Y} - \mathbf{Y}^d = \mathbf{T}\,\mathbf{R} - \mathbf{Y}^d \tag{5}$$

where $\mathbf{Y}^d$ is the reference trajectory vector. The optimal tracking problem can be expressed using (5) as

$$\mathbf{R}^* = \arg\min_{\mathbf{R}} J(\mathbf{R}) = E\{(1/N) \cdot \|\mathbf{E}(\mathbf{R})\|_2^2\}\,. \tag{6}$$

Equation (5) shows that knowledge on $\mathbf{T}$ would provide the optimal solution which makes the tracking error zero, i.e., $\mathbf{R} = \mathbf{T}^{-1}\mathbf{Y}^d$. However, $\mathbf{T}$ can be ill-conditioned and it is always subject to measurement errors; therefore $\mathbf{T}^{-1}$ cannot be used. A solution to the ILC-based iterative estimation of $\mathbf{T}$ is given in [8]. The optimization problem (6) can also be solved over the controller parameters using IFT [9].

III. Model-Free Estimation of Gradient

An iterative approach to solve (6) in the deterministic case is based on

$$\mathbf{R}_{j+1} = \mathbf{R}_j - \gamma_j \left.\frac{\partial J}{\partial \mathbf{R}}\right|_{\mathbf{R}=\mathbf{R}_j}. \tag{7}$$

As suggested in [6] and [7], the combination of (6) and (7) suggests an optimization approach to an ILC problem. A different approach, based on the results given in [5], will be presented as follows. The OF (6) is quadratic with respect to the vector $\mathbf{R}$ and the gradient of $J(\mathbf{R})$ in the deterministic case at each iteration j is expressed as

$$\left.\frac{\partial J}{\partial \mathbf{R}}\right|_{\mathbf{R}=\mathbf{R}_j} = 2\mathbf{T}^T\mathbf{E}_j\,. \tag{8}$$

Equation (8) suggests that the gradient information can be obtained either by an experimentally measured $\mathbf{T}$ or by using a special gradient experiment at each iteration. The model-free gradient experiments that lead to the vector $\mathbf{T}^T\mathbf{E}_j$ in the linear case are presented in [5].

The changes in the reference trajectory are of iterative nature and they are carried out in the vicinity of the nominal trajectory at the current iteration. The linearity assumption and operation can therefore be justified in this case. In order to allow for near-nominal experiment regimes to be used with linear systems and moreover to extend the applicability of the IRIT algorithm to nonlinear systems, a perturbation-based approach is proposed in order to obtain the gradient information near the nominal trajectory. The idea stems from [9]. The algorithm consists of the following steps:

Step A. Record the tracking error at the current iteration in the vector $\mathbf{E}_j$.

DOI: 10.1109/ECC.2014.6862222

Step B. Define the reversed vector $rev(\mathbf{E}_j)$

$$\begin{aligned} rev(\mathbf{E}_j) &= rev([e_j^t(0) \quad \dots \quad e_j^t(N-n-1)]^T) \\ &= [e_j^t(N-n-1) \quad \dots \quad e_j^t(0)]^T. \end{aligned} \tag{9}$$

Step C. Apply $\mathbf{R}_j + \mu \times rev(\mathbf{E}_j)$ as a reference input to the CS and obtain the output vector $\mathbf{Y}_G = \mathbf{T}(\mathbf{R}_j + \mu \times rev(\mathbf{E}_j))$, where the subscript G indicates the gradient. The scalar parameter μ is chosen such that the perturbation term $\mu \times rev(\mathbf{E}_j)$ produces only a small deviation around the nominal $\mathbf{R}_j$.

Step D. Since $\mathbf{Y}_j = \mathbf{T}\,\mathbf{R}_j$ is known from the nominal experiment, obtain $\mathbf{T}^T\mathbf{E}_j$ as

$$\mathbf{T}^T\mathbf{E}_j = (1/\mu) \times rev(\mathbf{Y}_G - \mathbf{Y}_j) \tag{10}$$

and apply (8) to get the gradient of the OF.

The selection of the parameter μ can be done automatically. The presented approach is at the core of the following developments using the LTI-based formulations.

The stochastic properties of ILC algorithms are treated in [6] and [7]. Two stochastic convergence conditions are imposed with respect to (7): the estimated gradient is unbiased, and the step size sequence $\{\gamma_j\}_{j\ge 0}$ converges to zero but not too fast.

IV. Iterative Reference Input Tuning Algorithm

A framework to deal with control signal saturation and rate constraints is next introduced. The lifted form representations allow the expression of a particular useful form of the optimization problem. Assuming the deterministic case, let $\mathbf{S}_{ur} \in \Re^{(N-m)\times(N-m)}$ be the lifted map that corresponds to the transfer function $S_{ur}(q^{-1}) = C(q^{-1})S(q^{-1})$, where $\Re$ is the set of real numbers. Using the notation m for the relative degree of $S_{ur}(q^{-1})$, $m \le n$, the lifted form representations are

$$\begin{aligned} \mathbf{U} &= [u(m) \quad u(m+1) \quad \dots \quad u(N-1)]^T, \\ \mathbf{R} &= [r(0) \quad r(1) \quad \dots \quad r(N-m-1)]^T, \\ \mathbf{S}_{ur} &= \begin{bmatrix} s_1 & 0 & \dots & 0 \\ s_2 & s_1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ s_{N-m} & s_{N-m-1} & \dots & s_1 \end{bmatrix}. \end{aligned} \tag{11}$$

The control signal vector can be expressed as $\mathbf{U} = \mathbf{S}_{ur}\mathbf{R}$, where $\mathbf{R} \in \Re^{(N-m)\times 1}$ is a vector of greater length than in (4), for which $\mathbf{R} \in \Re^{(N-n)\times 1}$. Therefore, a truncation of $\mathbf{S}_{ur}$ corresponding to the leading principal minor of size $N-n$ is considered such that $\mathbf{S}_{ur} \in \Re^{(N-n)\times(N-n)}$. This is because we want the same $\mathbf{R}$ of size $N-n$ to be tuned and this in turn will allow only $N-n$ (out of $N-m$) constraints imposed to $\mathbf{U}$ to be shown as follows, and the affine constraint $\mathbf{U}_{\min} \le \mathbf{U}(\mathbf{R}) \le \mathbf{U}_{\max}$ is imposed on $\mathbf{R}$. So even though we could benefit from the dimensionality of the map $\mathbf{S}_{ur}$, we choose only the appropriate size in order to tune the initial $\mathbf{R}$ from (4). The control signal vector can be expressed next as $\mathbf{U}(\mathbf{R}) = \mathbf{S}_{ur}\mathbf{R}$, where $\mathbf{S}_{ur}\mathbf{R} \in \Re^{(N-n)\times(N-n)}$, and the constraints hold for $2(N-n)$ lower and upper bounds.

Using (11), the control signal rate sequence $\Delta u(k)$ can be expressed in the lifted form

$$\begin{aligned} \Delta\mathbf{U} &= [\Delta u(1) \quad \Delta u(2) \quad \dots \quad \Delta u(N-n)]^T \\ &= [u(m) - 0 \quad u(m+1) - u(m) \quad \dots \\ &\quad u(m+N-n-1) - u(m+N-n-2)]^T \\ &= [s_1 r(0) \quad s_2 r(0) + s_1 r(1) - s_1 r(0) \quad \dots \\ &\quad s_{N-n} r(0) + \dots + s_1 r(N-n-1) - s_{N-n-1} r(0) \\ &\quad \dots - s_1 r(N-n-2)]^T = \begin{bmatrix} s_1 & s_2 & s_3 & \dots & s_{N-n} \\ 0 & s_1 & s_2 & \dots & s_{N-n-1} \\ 0 & 0 & s_1 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & s_1 \end{bmatrix}^T \cdot \mathbf{R} \\ &\quad - \begin{bmatrix} 0 & s_1 & s_2 & \dots & s_{N-n-1} \\ 0 & 0 & s_1 & \dots & s_{N-n-2} \\ 0 & 0 & 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T \cdot \mathbf{R} = \mathbf{S}_{\Delta ur} \cdot \mathbf{R}. \end{aligned} \tag{12}$$

As shown in (12), the control signal rate vector can be expressed as $\Delta\mathbf{U}(\mathbf{R}) = \mathbf{S}_{\Delta ur}\mathbf{R}$, and these constraints also hold for $2(N-n)$ lower and upper bounds.

Using the following notations for the vectors of lower and upper bounds:

$$\begin{aligned} &\mathbf{U}_{\min} = [u_{\min}^1 \dots u_{\min}^{N-n}]^T, \mathbf{U}_{\max} = [u_{\max}^1 \dots u_{\max}^{N-n}]^T, \\ &\Delta\mathbf{U}_{\min} = [\Delta u_{\min}^1 \dots \Delta u_{\min}^{N-n}]^T, \Delta\mathbf{U}_{\max} = [\Delta u_{\max}^1 \dots \Delta u_{\max}^{N-n}]^T, \end{aligned} \tag{13}$$

the inequality constraints are expressed as $\mathbf{U}_{\min} \le \mathbf{U}(\mathbf{R}) \le \mathbf{U}_{\max}$, $\Delta\mathbf{U}_{\min} \le \Delta\mathbf{U}(\mathbf{R}) \le \Delta\mathbf{U}_{\max}$, and the optimization problem which ensures the reference trajectory tracking with control signal constraints and with control signal rate constraints is expressed as

$$\begin{aligned} &\mathbf{R}^* = \arg\min_{\mathbf{R}} (1/N)(\mathbf{R}^T\,\mathbf{Q}\,\mathbf{R} + \mathbf{q}\,\mathbf{R} + \alpha), \\ &\text{subject to } \tilde{\mathbf{S}}\,\mathbf{R} \le \tilde{\mathbf{U}} \text{ and to } \tilde{\mathbf{S}}_\Delta\,\mathbf{R} \le \Delta\tilde{\mathbf{U}}, \\ &\tilde{\mathbf{S}} = [\mathbf{S}_{ur}^T \quad -\mathbf{S}_{ur}^T]^T \in \Re^{2(N-n)\times(N-n)}, \\ &\tilde{\mathbf{S}}_\Delta = [\mathbf{S}_{\Delta ur}^T \quad -\mathbf{S}_{\Delta ur}^T]^T \in \Re^{2(N-n)\times(N-n)} \\ &\tilde{\mathbf{U}} = [\mathbf{U}_{\max}^T \quad -\mathbf{U}_{\min}^T]^T \in \Re^{2(N-n)\times 1}, \\ &\Delta\tilde{\mathbf{U}} = [\Delta\,\mathbf{U}_{\max}^T \quad -\Delta\mathbf{U}_{\min}^T]^T \in \Re^{2(N-n)\times 1}, \end{aligned} \tag{14}$$

DOI: 10.1109/ECC.2014.6862222

where $\mathbf{Q} = \mathbf{T}^T\mathbf{T}$ is a positive semi-definite matrix, $\mathbf{M} = -\mathbf{Y}^d$, $\mathbf{q} = 2\,\mathbf{M}\,\mathbf{T}^T$, and $\alpha = \mathbf{M}^T\mathbf{M}$.

A solver for this type of problems in the deterministic case is the IPB algorithm [5]. The logarithmic barrier penalty function used in the IPB algorithm grows unbounded as the constraints are violated and in the stochastic framework this is always the case. A solution to this problem uses quadratic penalty functions [10]. We propose the following augmented OF that accounts for inequality constraints concerning the control signal saturation and the control signal rate:

$$\widetilde{J}_{p_j}(\mathbf{R}) = J(\mathbf{R}) + p_j[\underbrace{\frac{1}{2}\sum_{h=1}^{c}\{[\max\{0,-(\underbrace{\widetilde{u}_h - \widetilde{\mathbf{s}}_h^T\mathbf{R}}_{q_h(\mathbf{R})})\}]^2\}}_{\phi(\mathbf{R})} \quad (15)$$

$$+\underbrace{\frac{1}{2}\sum_{h=1}^{c}\{[\max\{0,-q_h(\mathbf{R})\}]^2\}}_{\Delta\phi(\mathbf{R})}],$$

where the positive and strictly increasing sequence of penalty parameters $\{p_j\}_{j\geq 0}$, $p_j \to \infty$, guarantees that the minimum of the sequence of augmented OFs $\{\widetilde{J}_{p_j}(\mathbf{R})\}_{j\geq 0}$ will converge to the solution to the constrained optimization problem (14), h, $h = 1...c$, is the constraint index, $q_h(\mathbf{R}) > 0$ is the h^{th} constraint, $\widetilde{u}_h$ is the h^{th} element of $\widetilde{\mathbf{U}}$, and $\widetilde{\mathbf{s}}_h^T$ is the h^{th} row of $\widetilde{\mathbf{S}}$. The optimization problem (14) is solved using a stochastic approximation algorithm which makes use of the experimentally obtained gradient of $\widetilde{J}_{p_j}(\mathbf{R})$.

The quadratic penalty functions $\phi(\mathbf{R})$ and $\Delta\phi(\mathbf{R})$ use the maximum function which in this case is non-differentiable only at zero. Given that $\phi(\mathbf{R})$ and $\Delta\phi(\mathbf{R})$ are Lipschitz and non-differentiable at a set of points of zero Lebesgue measure, the algorithm visits the zero-measure set with probability zero when a normal distribution for the noise is assumed [10]. Therefore, using the fact that

$$\frac{\partial[\max\{0,-q_h(\mathbf{R})\}]^2}{\partial r(i)} = -2\max\{0,-q_h(\mathbf{R})\}\frac{\partial q_h(\mathbf{R})}{\partial r(i)}, \quad (16)$$

and the gradients of $\phi(\mathbf{R})$ and $\Delta\phi(\mathbf{R})$ with respect to $\mathbf{R}$ taken from the appendix (Eqs. (21) and (23), respectively), the expression of the gradient of the OF (14) at the current iteration j is

$$\left.\frac{\partial\widetilde{J}(\mathbf{R})}{\partial\mathbf{R}}\right|_{\mathbf{R}=\mathbf{R}_j} = 2\,\mathbf{T}^T\mathbf{E}_j + p_j\{\mathbf{S}_{ur}^T\underbrace{[\zeta(\mathbf{R}_j) + \Delta\zeta(\mathbf{R}_j) - \Delta\bar{\zeta}(\mathbf{R}_j)]}_{\psi(\mathbf{R}_j)}\}, \quad (17)$$

where $\mathbf{R} \in \Re^{(N-n)\times 1}$, $\zeta(\mathbf{R}_j)$ is defined in (21) as $\zeta(\mathbf{R})$, $\Delta\zeta(\mathbf{R}_j)$ is defined in (23) as $\Delta\zeta(\mathbf{R})$, and $\Delta\bar{\zeta}(\mathbf{R}_j)$ is a one step ahead vector of dimension $N-n$:

$$\Delta\bar{\zeta}(\mathbf{R}_j) = [\Delta\zeta(\mathbf{R},2) \quad ... \quad \Delta\zeta(\mathbf{R},N-n) \quad 0]^T. \quad (18)$$

The matrix $\mathbf{M}_1$ in (23) is exactly the map $\mathbf{S}_{ur}^T$, and $\mathbf{M}_2$ is formed using the impulse response coefficients from $\mathbf{S}_{ur}^T$.

Therefore, the term $(\mathbf{M}_1 - \mathbf{M}_2)\Delta\zeta(\mathbf{R})$ can be obtained in one gradient experiment described as follows. It can be easily shown that $\mathbf{M}_1\Delta\bar{\zeta}(\mathbf{R}) = \mathbf{M}_2\Delta\zeta(\mathbf{R})$. Hence, a gradient scheme is used with $rev(\Delta\zeta - \Delta\bar{\zeta})$ injected as the reference input to the CS. The control signal in this experiment is recorded and then reversed. The $N-n$ samples of this vector are exactly $(\mathbf{M}_1 - \mathbf{M}_2)\Delta\zeta(\mathbf{R})$. This approach takes advantage of the dimensionality of the map $\mathbf{S}_{ur}^T$.

V. Case Study and Discussion of Simulation Results

The case study deals with the angular positioning of the vertical motion of a twin-rotor aero-dynamical system experimental setup with the process model and parameters given in [11]. The control signal is the PWM duty-cycle corresponding to the input voltage range of the DC motor and the output and $\alpha_v\,(rad) = y$ is the process output.

The nonlinear process model is not used in the tuning process except for obtaining an initial controller. A discrete-time linear PI controller with the transfer function $H(q^{-1}) = (0.12 + 0.012q^{-1})/(1-q^{-1})$ is considered. Similar PI controllers eventually combined with additional functionalities are treated in [12]–[14].

The reference trajectory is prescribed in terms of the unit step response of a second-order normalized reference model with the transfer function $\omega_n^2/(s^2 + 2\zeta\omega_n s + \omega_n^2)$ with $\omega_n = 0.5$ rad/s and $\zeta = 0.7$. The sampling period is $T_s = 0.1$ s for experiment length of $N = 200$ samples. The relative degree of $T(q^{-1})$ is $n = 1$ and the relative degree of $S_{ur}(q^{-1})$ is $m = 0$. The initial reference signal is chosen such as the closed-loop step response is very different from the targeted reference trajectory. The value of the initial reference is $r(k) = (1/200)\cdot k \cdot T_s$.

A simulated case study corresponding to the optimization with control signal saturation and rate constraints is first shown; the two inequality constraints are $-0.05 \leq u \leq 0.14$ and $-0.04 \leq \Delta u \leq 0.04$. A normally distributed zero-mean white noise with variance 10^{-10} was introduced.

Even if the noise is added, the algorithm is applied as in the deterministic case as follows. The sequence of penalty parameters in (15) was set to a constant value $p_j = 45$. Two constant values of the step-scaling parameter were used for the gradient descent. Prior to any constraint violation the step scaling is $\gamma = 150$, and it is next set to $\gamma = 50$. 200 samples of the reference input are subject to optimization and a total of 796 constraints were used: 398 for control signal saturation and 398 for control signal rate saturation.

Fig. 1 shows strong reference input changes even after the first iteration. The initial reference input is the ramp ending at 0.1 rad at 20 s. Since both the reference trajectory and the initial output response start with zero, the error signal between the two at the first sample time is zero. When this error signal is reversed, scaled and added to the nominal

DOI: 10.1109/ECC.2014.6862222

reference in the gradient experiment, it does not produce a disturbance at the final sample of the reference. Because of this lack of excitation, the initial reference input has to end at the same final value as the one of the reference trajectory. With this respect, nonzero initial conditions for the CS output actually help. The final reference input shows a constant value for the first 1.5 s, which prevents constraint violation of the control signal rate illustrated in Fig. 2.

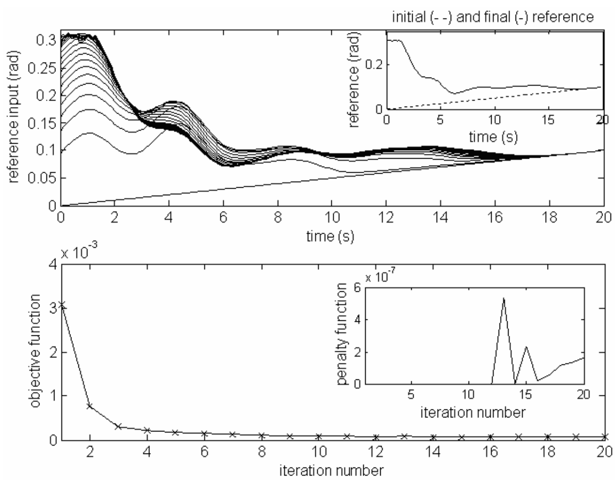

Figure 1. Simulation results expressed as reference input versus time as the learning converges and as OF versus iteration number.

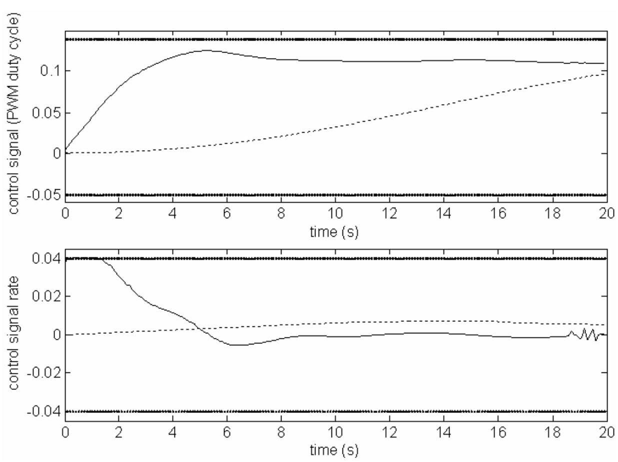

Figure 2. Simulation results expressed as control signal responses: initial (dashed), and final (solid), constant upper and lower bounds (dotted).

The evolution of the OF presented in Fig. 1 shows a fast convergence considering the search space of dimension 200. The penalty function indicates that the constraints are firstly violated at the 12th iteration of the algorithm.

Fig. 2 highlights how the control signal and the control signal rate evolve from the initial responses to the final ones. When control signal rate constraints are violated first the tuning is driven in the direction of shaping the reference input such that the constraints are fulfilled and the reference tracking is ensured. Fig. 2 also shows strong improvements.

The noise affects the numeric differentiation that occurs when the gradient is estimated (in trial domain) and also when the sample derivative of the control signal rate is calculated (in time domain). Several simulations with significantly increased noise intensity show that the reference tuning still works.

The same significant performance improvement against the initial CS with PI controller and compared to [12]–[14] is shown in Fig. 3, where the final response after 20 iterations follows the reference trajectory in an acceptable manner given the constraints. Even if the minimum is not reached the performance improvements are considerable.

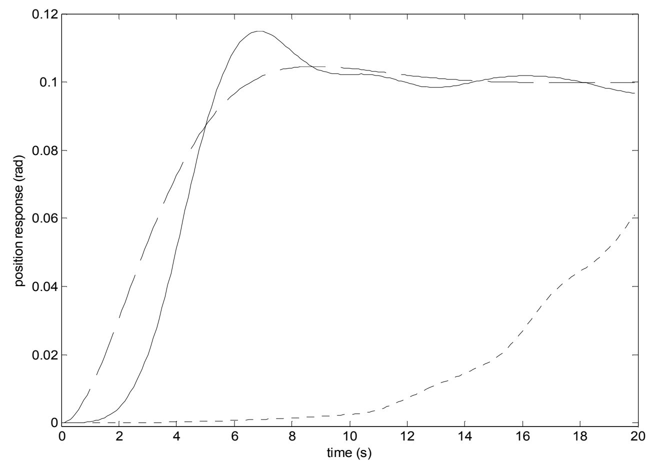

Figure 3. Simulation results expressed as position response: initial (dotted), final response after optimization (solid) and reference trajectory (dashed).

VI. CONCLUSION

This paper has proposed a data-driven algorithm that solves an optimal control problem in order to ensure the reference trajectory tracking by few experiments conducted on the real-world control system in a representative mechatronics application. Our IRIT algorithm is also advantageous as it works for smooth nonlinear systems.

Not only the gradient experiment is conducted once per iteration in order to obtain the gradient information but it is also performed in the vicinity of the nominal trajectory. This is an advantage since the abnormal experiment regimes are not required for the CS.

Our algorithm can be generalized by considering other data-driven optimization approaches to controller design and tuning [2]–[4], with several objective functions [14], [15], combined with ILC to optimize the reference input sequence.

APPENDIX

The quadratic penalty function $\phi(\mathbf{R})$ can be expressed as:

$$\begin{aligned}
\phi(\mathbf{R}) = \frac{1}{2}\{&[\max\{0,-(\tilde{u}_1 - s_1 r(0))\}]^2 \\
&+[\max\{0,-(\tilde{u}_2 - s_2 r(0) - s_1 r(1))\}]^2 + \ldots \\
&+[\max\{0,-(\tilde{u}_{N-n} - s_{N-n} r(0) - \ldots - s_1 r(N-n-1))\}]^2 \\
&+[\max\{0,-(\tilde{u}_{N-n+1} + s_1 r(0))\}]^2 \\
&+[\max\{0,-(\tilde{u}_{N-n+2} + s_2 r(0) + s_1 r(1))\}]^2 + \ldots \\
&+[\max\{0,-(\tilde{u}_{2(N-n)} + s_{N-n} r(0) + \ldots + s_1 r(N-n-1))\}]^2\}.
\end{aligned} \quad (19)$$

The gradient with respect to $r(0)$ is

DOI: 10.1109/ECC.2014.6862222

$$\frac{\partial\phi(\mathbf{R})}{r(0)} = s_1 \max\{0,-(\widetilde{u}_1 - s_1 r(0))\}$$
$$+ s_2 \max\{0,-(\widetilde{u}_2 - s_2 r(0) - s_1 r(1))\} + \ldots$$
$$+ s_{N-n} \max\{0,-(\widetilde{u}_{N-n} - s_{N-n} r(0) - \ldots \tag{20}$$
$$- s_1 r(N-n-1))\}]^2 - s_1 \max\{0,-(\widetilde{u}_{N-n+1} + s_1 r(0))\}$$
$$- s_2 \max\{0,-(\widetilde{u}_{N-n+2} + s_2 r(0) + s_1 r(1))\} - \ldots$$
$$- s_{N-n} \max\{0,-(\widetilde{u}_{2(N-n)} + s_{N-n} r(0) + \ldots + s_1 r(N-n-1))\}\}.$$

Using relationships that are similar to (20) for the other components of **R**, the matrix form of the gradient of $\phi(\mathbf{R})$ with respect to **R** is

$$\frac{\partial\phi(\mathbf{R})}{\partial\mathbf{R}} = \begin{bmatrix} s_1 & s_2 & \ldots & s_{N-n} \\ 0 & s_1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & s_1 \end{bmatrix} \cdot (\boldsymbol{\varepsilon}_G^1(\mathbf{R}) - \boldsymbol{\varepsilon}_G^2(\mathbf{R})) = \mathbf{S}_{ur}^T \cdot \boldsymbol{\zeta}(\mathbf{R}),$$
$$\boldsymbol{\zeta}(\mathbf{R}) = \boldsymbol{\varepsilon}_G^1(\mathbf{R}) - \boldsymbol{\varepsilon}_G^2(\mathbf{R}), \tag{21}$$
$$\boldsymbol{\varepsilon}_G^1(\mathbf{R}) = [\max\{0,-q_1(\mathbf{R})\} \quad \ldots \quad \max\{0,-q_{N-n}(\mathbf{R})\}]^T,$$
$$\boldsymbol{\varepsilon}_G^2(\mathbf{R}) = [\max\{0,-q_{N-n+1}(\mathbf{R})\} \quad \ldots \max\{0,-q_{2(N-n)}(\mathbf{R})\}]^T.$$

Using (15), $\Delta\phi(\mathbf{R})$ can be expressed as:

$$\Delta\phi(\mathbf{R}) = \frac{1}{2}\{(\max\{0,-\Delta u_{\max}^1 + s_1 r(0)\})^2$$
$$+ (\max\{0,-\Delta u_{\max}^2 + s_2 r(0) + s_1 r(1) - s_1 r(0)\})^2 + \ldots$$
$$+ (\max\{0,-\Delta u_{\max}^{N-n} + s_{N-n} r(0) + \ldots s_1 r(N-n-1) \tag{22}$$
$$- s_{N-n-1} r(0) - \ldots - s_1 r(N-n-2)\})^2$$
$$+ (\max\{0,\Delta u_{\min}^1 - s_1 r(0)\})^2$$
$$+ (\max\{0,\Delta u_{\min}^2 - s_2 r(0) - s_1 r(1) + s_1 r(0)\})^2 + \ldots$$
$$+ (\max\{0,\Delta u_{\min}^{N-n} - s_{N-n} r(0) - \ldots - s_1 r(N-n-1)$$
$$+ s_{N-n-1} r(0) + \ldots + s_1 r(N-n-2)\})^2\}.$$

Using (16) in (22), the gradient of $\Delta\phi(\mathbf{R})$ with respect to **R** will obtain the expression

$$\frac{\partial\Delta\phi(\mathbf{R})}{\partial\mathbf{R}} = \begin{bmatrix} s_1 & s_2 & s_3 & \ldots & s_{N-n} \\ 0 & s_1 & s_2 & \ldots & s_{N-n-1} \\ 0 & 0 & s_1 & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & s_1 \end{bmatrix} \cdot \Delta\boldsymbol{\zeta}(\mathbf{R}) \tag{23}$$
$$- \begin{bmatrix} 0 & s_1 & s_2 & \ldots & s_{N-n-1} \\ 0 & 0 & s_1 & \ldots & s_{N-n-2} \\ 0 & 0 & 0 & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot \Delta\boldsymbol{\zeta}(\mathbf{R}) = (\mathbf{M}_1 - \mathbf{M}_2)\Delta\boldsymbol{\zeta}(\mathbf{R}),$$

where

$$\Delta\boldsymbol{\zeta}(\mathbf{R}) = \Delta\boldsymbol{\varepsilon}_G^1(\mathbf{R}) - \Delta\boldsymbol{\varepsilon}_G^2(\mathbf{R}) =$$
$$= [\Delta\zeta(\mathbf{R},1) \quad \ldots \quad \Delta\zeta(\mathbf{R}, N-n-1) \quad \Delta\zeta(\mathbf{R}, N-n)]^T,$$
$$\Delta\boldsymbol{\varepsilon}_G^1 = \begin{bmatrix} \max\{0,-\Delta u_{\max}^1 + s_1 r(0)\} \\ \ldots \\ \max\{0,-\Delta u_{\max}^{N-n} + s_{N-n} r(0) + \ldots + s_1 r(N-n-1) \\ - s_{N-n-1} r(0) - \ldots - s_1 r(N-n-2)\} \end{bmatrix}, \tag{24}$$
$$\Delta\boldsymbol{\varepsilon}_G^2 = \begin{bmatrix} \max\{0,\Delta u_{\min}^1 - s_1 r(0)\} \\ \ldots \\ \max\{0,\Delta u_{\min}^{N-n} - s_{N-n} r(0) - \ldots - s_1 r(N-n-1) \\ + s_{N-n-1} r(0) + \ldots + s_1 r(N-n-2)\} \end{bmatrix}.$$

REFERENCES

[1] H. Hjalmarsson, "Iterative feedback tuning - an overview," *Int. J. Adapt. Control Signal Process.*, vol. 16, pp. 373–395, May 2002.

[2] J. C. Spall and J. A. Cristion, "Model-free control of nonlinear stochastic systems with discrete-time measurements," *IEEE Trans. Autom. Control*, vol. 43, pp. 1198–1210, Sep. 1998.

[3] S. Andary, A. Chemori, M. Benoit, and J. Sallantin, "A dual model-free control of underactuated mechanical systems, application to the inertia wheel inverted pendulum," in *Proc. 2012 American Control Conference*, Montréal, QC, Canada, 2012, pp. 1029–1034.

[4] S. Formentin, P. De Filippi, M. Corno, M. Tanelli, and S. M. Savaresi, "Data-driven design of braking control systems," *IEEE Trans. Contr. Syst. Technol.*, vol. 21, pp. 186–193, Jan. 2013.

[5] M.-B. Radac, R.-E. Precup, E. M. Petriu, S. Preitl, and C.-A. Dragos, "Data-driven reference trajectory tracking algorithm and experimental validation," *IEEE Trans. Ind. Informat.*, vol. 9, pp. 2327–2336, Nov. 2013.

[6] S. Gunnarsson and M. Norrlöf, "On the design of ILC algorithms using optimization," *Automatica*, vol. 37, pp. 2011–2016, Dec. 2001.

[7] M. Butcher, A. Karimi, and R. Longchamp, "Iterative learning control based on stochastic approximation," in *Proc. 17th IFAC World Congress*, Seoul, Korea, 2008, pp. 1478–1483.

[8] P. Janseens, G. Pipeleers, and J. Swevers, "A data-driven constrained norm-optimal iterative learning control framework for LTI systems," *IEEE Trans. Contr. Syst. Technol.*, vol. 21, pp. 546–551, Mar. 2013.

[9] J. Sjöberg, F. De Bruyne, M. Agarwal, B. D. O. Anderson, M. Gevers, F. J. Kraus, and N. Linard, "Iterative controller optimization for nonlinear systems," *Control Eng. Pract.*, vol. 11, pp. 1079–1086, Sep. 2003.

[10] I.-J. Wang and J. C. Spall, "Stochastic optimization with inequality constraints using simultaneous perturbations and penalty functions," *Int. J. Control*, vol. 81, pp. 1232–1238, Aug. 2008.

[11] M.-B. Radac, R.-C. Roman, R.-E. Precup, E. M. Petriu, C.-A. Dragos, and S. Preitl, "Data-based tuning of linear controllers for MIMO twin rotor systems," in *Proc. IEEE Region 8 EuroCon 2013 Conference*, Zagreb, Croatia, 2013, pp. 1915–1920.

[12] S. F. Toha and M. O. Tokhi, "PID and inverse-model-based control of a twin rotor system," *Robotica*, vol. 29, pp. 929–938, Oct. 2011.

[13] A. P. S. Ramalakshmi and P. S. Manoharan, "Non-linear modeling and PID control of twin rotor MIMO system," in *Proc. 2012 IEEE International Conference on Advanced Communication Control and Computing Technologies*, Ramanathapuram, India, 2012, pp. 366–369.

[14] D.-W. Gu, P. H. Petkov, and M. M. Konstantinov, *Robust Control Design with MATLAB®, 2nd ed.* London: Springer-Verlag, 2013.

[15] P. V. Pakshin, K. Gałkowski, and E. Rogers, "Linear-quadratic parametrization of stabilizing controls in discrete-time 2D systems," *Autom. Remote Control*, vol. 72, pp. 2364–2378, Nov. 2011.