Data-Driven Reference Trajectory Tracking Algorithm and Experimental Validation

Mircea-Bogdan Rădac, *Member, IEEE*, Radu-Emil Precup, *Senior Member, IEEE*, Emil M. Petriu, *Fellow, IEEE*, Stefan Preitl, *Senior Member, IEEE*, and Claudia-Adina Dragoş

***Abstract*—This paper proposes a data-driven algorithm that solves a reference trajectory tracking problem defined as an optimization problem. The new data-driven reference trajectory tracking algorithm (DDRTTA) solves the optimization problem in the framework of Iterative Learning Control (ILC). The DDRTTA updates the reference input sequence using an experiment-based approach which accounts for operational constraints and employs an interior point barrier algorithm. Therefore the DDRTTA combines the advantages of data-driven control and ILC. A case study which deals with the angular position control of a nonlinear servo system is included to validate the DDRTTA by experimental and simulation results.**

***Index Terms*—Data-driven reference trajectory tracking, experimental results, Iterative Feedback Tuning, Iterative Learning Control, lifted form representation.**

I. Introduction

DATA-DRIVEN optimization techniques for controller design offer the improvement of control system (CS) performance using no a priori model information about the controlled process or little such information [1]–[6]. The CS performance improvement offered by data-driven techniques is achieved by simple specifications and easily interpretable performance indices. These indices are usually specified in the time domain (e.g., rise time, overshoot, settling time), and they are aggregated in integral-type or sum-type objective functions (o.f.s) such as the Linear Quadratic Gaussian (LQG) ones. The minimization of the o.f.s in constrained optimization problems can fulfill different objectives 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) [7], Correlation-based Tuning [8], Frequency Domain Tuning [9], Iterative Regression Tuning [10], Simultaneous Perturbation Stochastic Approximation [11], [12], pulse response based control [13], Markov data-based LQG control [14], data-driven or data-based predictive control [15]–[17], LQ data-driven control [18], and the most popular non-iterative technique is Virtual Reference Feedback Tuning (VRFT) [19] These techniques use various approaches to ensure model-free controller tuning. However, the tuning for reference trajectory tracking does not guarantee robust stability or robust performance. Some recent data-driven control results ensure robust stability/performance while still preserving the model-free property; these results try to avoid the direct process identification and to infer the results from data or from easy-to-obtain non-parametric CS models such as the frequency response functions [9], [20].

The reference trajectory tracking can be considered as a reference input design over an initial CS with a priori tuned controllers for stability and disturbance rejection. The reference trajectory tracking is thus defined as an open-loop optimal control problem. The iterative experiment-based solving of this optimization problem makes use of IFT, where a gradient search algorithm is used and the gradient information is obtained experimentally without involving any process knowledge. With this regard the iterative controller parameter update laws are replaced by reference input sequence update laws and the Iterative Learning Control (ILC) framework [21], [22] is applied. General approaches to ILC-based solving of optimal control problems are reported in [23], [24], with time and frequency domain convergence analysis conducted in [25], stochastic approximation treated in [26], and output tracking given in [27].

Several techniques dealing with data-driven modeling and control are neural networks and fuzzy models [28]–[40]. Some data-driven techniques used in process control, associated with appropriately defined optimization problems [41]–[45], consists of Reinforcement Learning (RL), approximate dynamic programming and model-free predictive control. The difference between RL and our approach is that RL [46]–[48] is unsupervised learning whereas our approach is a form of

Manuscript received February 28, 2012; revised June 22, 2012; accepted September 17, 2012. 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 of the NSERC of Canada.

M.-B. Rădac, R.-E. Precup (corresponding author), S. Preitl, and C.-A. Dragoş are with the Department of Automation and Applied Informatics, "Politehnica" University of Timisoara, 300223 Timisoara, Romania (e-mail: mircea.radac@aut.upt.ro; radu.precup@aut.upt.ro; stefan.preitl@aut.upt.ro; claudia.dragos@aut.upt.ro).

E. M. Petriu is with the School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: petriu@eecs.uottawa.ca).

supervised learning, but the similarities between these two approaches are:

- The interaction with the unknown environment for collecting data. In our approach we aim to reduce the information about the process which is subject to control. Hence RL and reference trajectory tracking are connected by their data-driven approaches.
- The dynamic programming underlying problem. At each time, a decision is made to maximize the reward. The reward in our setting is to minimize the o.f. defined as the difference between the controlled process output and the prescribed reference input trajectory. We do not solve the problem using dynamic programming techniques, but we transform it into a static optimization problem and we solve it offline. This is carried out after trials based on data collection and o.f. evaluation. In dealing with the optimization problem we have to cope with the stochasticity of the real world affecting the collected data. Our approach is a direct policy search using stochastic optimization.

The main contribution of this paper with respect to the state-of-the-art is a new data-driven reference trajectory tracking algorithm (DDRTTA) which iteratively solves an optimal control problem that ensures reference trajectory tracking. The DDRTTA employs an entirely experimental-driven interior point barrier (IPB) algorithm which allows for ILC-based optimization accounting for several operational constraints. The importance and advantages of this new idea with respect to the previously analyzed literature are:

- The DDRTTA uses experiments conducted on the real-world CS; therefore it can compensate for process nonlinearities and uncertainties.
- The DDRTTA is characterized by a reduced number of experiments which ensures cost-effective implementations.

Therefore the data-driven optimization algorithm for controller design suggested in this paper is general and applicable to many complex industrial systems.

The paper is structured as follows. The next section presents the problem setting related to the reference trajectory tracking problem solved by data-driven iterative optimization techniques. Section III considers the experiment-based model-free estimation of the gradient of the o.f. and offers convergence analysis details. Section IV gives the model-free constrained optimal control problem and formulates the DDRTTA. Section V discusses the DDRTTA application on a case study dealing with the position control of a nonlinear servo system and validates the new DDRTTA. The conclusions are outlined in Section VI.

II. Problem Setting

The CS is dedicated to reference trajectory tracking by means of a reference model which specifies the performance indices. The process is characterized by the discrete time Linear Time-Invariant (LTI) Single Input-Single Output (SISO) model

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

where k is the discrete time argument, $y(k)$ is the process output, $r(k)$ is the reference input, $v(k)$ is the zero-mean stationary and bounded stochastic disturbance acting on the process output and accounting for various types of load or measurement disturbances, $S(\boldsymbol{\rho}, q^{-1}) = 1/[1 + P(q^{-1}) C(\boldsymbol{\rho}, q^{-1})]$ and $T(\boldsymbol{\rho}, q^{-1}) = 1 - S(\boldsymbol{\rho}, q^{-1})$ are the sensitivity function and the complementary sensitivity function, respectively, $P(q^{-1})$ is the process transfer function, $C(\boldsymbol{\rho}, q^{-1})$ is the controller transfer function, parameterized by the parameter vector $\boldsymbol{\rho}$ which contains the tuning parameters of the controller, and q^{-1} is the one step delay operator. $\boldsymbol{\rho}$ will be omitted as follows in some equations for the sake of simplicity.

The reference trajectory $y^d(k)$ can be generated by a reference model. The control signal $u(k)$ is not explicit in (1), but it may represent interest in control when constraints are required and the control effort should be manipulated. The objective of trajectory tracking is the minimization of the o.f. J expressed as the expected value for the Euclidean norm of the tracking error over the finite time horizon N:

$$J(\boldsymbol{\rho}, r) = E\{\| e^t \|_2^2 = (1/N) \sum_{k=0}^{N} [y(\boldsymbol{\rho}, r, k) - y^d(k)]^2\}\,, \tag{2}$$

where $e^t(\boldsymbol{\rho}, k) = y(\boldsymbol{\rho}, r, k) - y^d(k)$ is the tracking error and operational constraints can be accounted for.

Using (1), the minimization of (2) can be carried out by modifying either one or both the design variables $\boldsymbol{\rho}$ and r which translates (2) to a controller tuning or to a reference input modification. As shown in [49], the modification of either of the design variables creates the same effect in $y(k)$ as it is pointed out in the first order Taylor series expansion around the nominal values of $\boldsymbol{\rho}$ and r, with the notations $\boldsymbol{\rho}_n$ and r_n, respectively:

$$\begin{aligned} y(\boldsymbol{\rho}, r) &= y(\boldsymbol{\rho}_n, r) + (\partial y / \partial \boldsymbol{\rho} |_{\boldsymbol{\rho}=\boldsymbol{\rho}_n})(\boldsymbol{\rho} - \boldsymbol{\rho}_n) + \text{h.o.t.} \\ &\approx y(\boldsymbol{\rho}, r_n) + (\partial y / \partial r |_{r=r_n})(r - r_n) + \text{h.o.t.}, \end{aligned} \tag{3}$$

where h.o.t. points out the higher order terms.

For the simultaneous tuning of $\boldsymbol{\rho}$ and r, Eq. (3) can be interpreted as a two-degree-of-freedom (2-DOF) structure tuning, where the feedback controller $C(\boldsymbol{\rho})$ is designed for disturbance input regulation and the feedforward filter $F(\boldsymbol{\rho})$ is designed for reference input tracking as illustrated in Fig. 1, with σ – the step reference input and e – the control error.

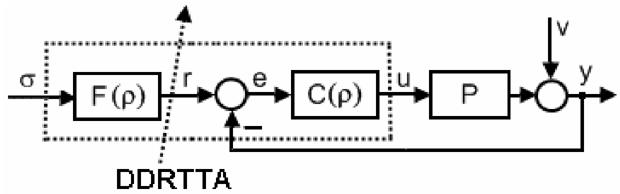

Fig. 1. 2-DOF control system structure for reference trajectory tracking.

As mentioned in Section I, the trajectory tracking can be formulated in terms of the following open-loop optimal control problem expressed as the static optimization problem

$$\begin{aligned} & r^* = \arg\min_r J(\boldsymbol{\rho}, r), \\ & \textit{subject to system dynamics } (1) \\ & \textit{and to some operational constraints.} \end{aligned} \tag{4}$$

Model information about the process is essential in order to solve the problem (4) analytically, but the discrepancies between the model and the reality could still compromise the controller design. IFT can be an alternative since the tuning is based on a gradient search algorithm for which the gradient information is obtained experimentally without any process knowledge. The following iterative update law can be employed and expressed as a steepest descent approach:

$$r_{j+1} = r_j - \gamma_j est\{\partial J / \partial r|_{r=r_j}\}, \tag{5}$$

where the subscript j indicates the iteration number, $j \in \mathbf{N}$, $\mathbf{N}$ is the set of natural numbers, the estimate of the gradient $est\{\partial J / \partial r|_{r=r_j}\}$ should be obtained experimentally, and γ_j is a positive step-scaling parameter. This approach basically integrates the search algorithm (5) into an ILC problem to apply the ILC framework [22]–[24]. 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 $\mathbf{Y} = \mathbf{T}\,\mathbf{R} + \mathbf{Y}_0$, with the matrices

$$\begin{aligned} & \mathbf{Y} = [y(n) \quad y(n+1) \quad \dots \quad y(N-1)]^T, \\ & \mathbf{R} = [r(0) \quad r(1) \quad \dots \quad r(N-n-1)]^T, \\ & \mathbf{Y}_0 = [y_{10} \quad y_{20} \quad \dots \quad y_{(N-n)0}]^T, \\ & \mathbf{T} = \begin{bmatrix} t_1 & 0 & \dots & 0 \\ t_2 & t_1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ t_{N-n} & t_{N-n-1} & \dots & t_1 \end{bmatrix}, \end{aligned} \tag{6}$$

$\mathbf{R}$ 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 $\mathbf{E} = \mathbf{Y} - \mathbf{Y}^d = \mathbf{T}\,\mathbf{R} - \mathbf{Y}^d$, where $\mathbf{Y}^d$ is the reference trajectory vector. 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.

III. Experiment-Based Model-Free Estimation of Gradient and Convergence Analysis

The vector form of (5) in the deterministic case is

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

As analyzed in Section I, the combination of (4) and (5) as an optimization approach to the ILC problem is treated in [24]–[27], and a new approach will be presented as follows. The o.f. (2) is quadratic with respect to the vector $\mathbf{R}$ and the gradient of J in the deterministic case at each iteration j is

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

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 choice of one or another one of these options depends on the measurement noise which affects the quality of $\mathbf{T}$ and on the possibility to conduct experiments on the CS.

The gradient experiments may affect the nominal operating regimes. To avoid this, gradient experiments are conducted in terms of steps A to D leading to the vector $\mathbf{T}^T\mathbf{E}_j$:

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

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 $rev(\mathbf{E}_j)$ as reference input to the CS and obtain the output vector $\mathbf{Y} = \mathbf{T}\,rev(\mathbf{E}_j)$.

Step D. Obtain $\mathbf{T}^T\mathbf{E}_j$ as

$$\mathbf{T}^T\mathbf{E}_j = rev(\mathbf{T}\,rev(\mathbf{E}_j)), \tag{10}$$

and use it to obtain the gradient in (8).

The four steps of this relatively simple model-free approach can be used to obtain the gradient information.

The convergence of the steepest descent approach is briefly studied as follows in the deterministic setting. Following (8), the tracking error vector at the next iteration is

$$\begin{aligned}\mathbf{E}_{j+1} &= \mathbf{T}\mathbf{R}_{j+1} - \mathbf{Y}^d = \mathbf{T}(\mathbf{R}_j - 2\gamma_j \mathbf{T}^T \mathbf{E}_j) - \mathbf{Y}^d \\ &= \mathbf{T}[\mathbf{T}^{-1}(\mathbf{E}_j + \mathbf{Y}^d) - 2\gamma_j \mathbf{T}^T \mathbf{E}_j] - \mathbf{Y}^d = (\mathbf{I} - 2\gamma_j \mathbf{T}\mathbf{T}^T)\mathbf{E}_j.\end{aligned} \tag{11}$$

The asymptotic convergence is ensured for the spectrum of the matrix $(\mathbf{I} - 2\gamma_j \mathbf{T}\mathbf{T}^T)$ in the unit disk. However, from a practical point of view, monotonic convergence is preferred in order to ensure a decrease of the tracking error at each iteration. This can be achieved by making the L_2-induced norm (the spectral norm) of $(\mathbf{I} - 2\gamma_j \mathbf{T}\mathbf{T}^T)$ less than unity to be able to express the convergence condition $\| E_{j+1} \|_2 \le \| E_j \|_2$. This condition together with (11) lead to another expression of the convergence condition $\bar{\sigma}(\mathbf{I} - 2\gamma_j \mathbf{T}\mathbf{T}^T) < 1$, where $\bar{\sigma}$ stands for the maximum singular value. Therefore, a suitable $\gamma_j > 0$ can be found at each iteration as an offline solution to the nonlinear optimization problem

$$\begin{aligned}&\gamma_j^* = \arg\min_{\gamma_j} \gamma_j, \\ &\textit{subject to } \bar{\sigma}(\mathbf{I} - 2\gamma_j \mathbf{T}\mathbf{T}^T) < 1.\end{aligned} \tag{12}$$

When $\mathbf{T}$ is obtained via informative experiments such as impulse response, uncertainties due to measurement errors have to be accounted for. Lifted form representations under multiplicative uncertainty can be used with this regard and expressed as $\mathbf{T} = \widetilde{\mathbf{T}}(\mathbf{I} + \boldsymbol{\Delta})$, where the bounds for $\boldsymbol{\Delta}$ can be estimated experimentally and can be included in (12).

The stochastic properties of ILC algorithms are studied in [25], [26]. Robbins-Monro's stochastic approximation algorithm can be employed to find the zero of the gradient of the o.f. The stochastic convergence conditions are:

- I. The estimated gradient of the o.f. is unbiased.
- II. The step size sequence $\{\gamma_j\}_{j \ge 0}$ converges to zero but not too fast.

Condition I. is fulfilled as in the gradient experiments scheme the recorded error and the output obtained with the error as input are trial-uncorrelated with respect to the noise. Condition II. is guaranteed for the following choice of γ_j:

$$\sum_{j=0}^{\infty} \gamma_j = \infty,\ \sum_{j=0}^{\infty} \gamma_j^2 < \infty,\ \gamma_j > 0\ \forall j \ge 0 \cdot \tag{13}$$

IV. Data-Driven Algorithm

The operational constraints regarding the saturation of actuators or the bounds on the process state variables are very important in the majority of complex industrial control systems applications. Different numerical algorithms can be employed to solve the optimization problem (4) for such systems. However, if the model-free approach proposed in the previous section is applied, this leads to a small number of iterations and to a small number of informative experiments that can affect the normal operating regimes.

The lifted form representations allow the expression of a particular form of the optimization problem which can be of interest. 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

$$\begin{aligned}&\mathbf{U} = [u(m) \quad u(m+1) \quad \ldots \quad u(N-1)]^T, \\ &\mathbf{R} = [r(0) \quad r(1) \quad \ldots \quad r(N-m-1)]^T, \\ &\mathbf{S}_{ur} = \begin{bmatrix} s_1 & 0 & \ldots & 0 \\ s_2 & s_1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ s_{N-m} & s_{N-m-1} & \ldots & s_1 \end{bmatrix}.\end{aligned} \tag{14}$$

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 (6), 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 (6). The o.f. $J(\mathbf{R})$ is quadratic with respect to $\mathbf{R}$:

$$\begin{aligned}J(\mathbf{R}) &= E\{(1/N)(\underbrace{\mathbf{T}\,\mathbf{R} - \mathbf{Y}^d}_{\mathbf{M}})^T (\mathbf{T}\,\mathbf{R} - \mathbf{Y}^d)\} \\ &= E\{(1/N)(\mathbf{R}^T\,\mathbf{Q}\,\mathbf{R} + \mathbf{q}\,\mathbf{R} + \alpha)\},\end{aligned} \tag{15}$$

with $\mathbf{Q} = \mathbf{T}^T\mathbf{T}$ – positive semi-definite, $\mathbf{q} = 2\,\mathbf{M}\,\mathbf{T}^T$ and $\alpha = \mathbf{M}^T\mathbf{M}$. The optimization problem which ensures reference trajectory tracking is expressed as

$$\begin{aligned}&\mathbf{R}^* = \arg\min_{\mathbf{R}} (\mathbf{R}^T\,\mathbf{Q}\,\mathbf{R} + \mathbf{q}\,\mathbf{R} + \alpha), \\ &\textit{subject to } \widetilde{\mathbf{S}}\,\mathbf{R} \le \widetilde{\mathbf{U}},\ \widetilde{\mathbf{S}} = [\mathbf{S}_{ur}^T \quad -\mathbf{S}_{ur}^T]^T \in \Re^{2(N-n)\times(N-n)}, \\ &\widetilde{\mathbf{U}} = [\mathbf{U}_{\max}^T \quad -\mathbf{U}_{\min}^T]^T \in \Re^{2(N-n)\times 1}.\end{aligned} \tag{16}$$

Efficient convex quadratic programming solvers can be applied to (16), such as the IPB algorithm [50]. The IPB algorithm will be adapted to this problem as follows in order to obtain the DDRTTA which solves (16). The idea is to transform (16) into the unconstrained optimization problem

$$\mathbf{R}^* = \arg\min_{\mathbf{R}} (\mathbf{R}^T\,\mathbf{Q}\,\mathbf{R} + \mathbf{q}\,\mathbf{R} + \alpha + \kappa\,\phi(\mathbf{R})), \tag{17}$$

with $\kappa > 0$, and $\phi(\mathbf{R}) = -\sum_{i=1}^{c} \log(\widetilde{u}_i - \widetilde{\mathbf{s}}_i^T \mathbf{R})$ is the logarithmic barrier function, where $\widetilde{u}_i$ is the i^{th} element of of $\widetilde{\mathbf{U}}$, $\widetilde{\mathbf{s}}_i^T$ is the i^{th} row of $\widetilde{\mathbf{S}}$, and $c = 2(N-n)$ is the number of constraints. The violation of constraints makes $\phi(\mathbf{R})$ and next the o.f. go to infinity.

The o.f. including the barrier function can first be evaluated in one experiment at the current reference input since the control signal can be recorded. An experiment-based informed search can next be applied to (17), and this requires the gradient of the o.f. with respect to **R**. The gradient of the first three terms of the new o.f. in (17) can be obtained by the informative experiments indicated in the previous section. The gradient of the remaining of the o.f. can also be obtained experimentally as shown in the sequel.

For an N length experiment, $\mathbf{R} \in \Re^{(N-n)\times 1}$, $2(N-n)$ constraints are available for a lower and upper bounded input as all constraints are applied to the first $N-n$ control signal samples starting with $u(m)$. Accepting that the optimization problem (17) and $\phi(\mathbf{R})$ implicitly have only $N-n$ constraints for the upper bound of the control signal, it can be shown by chain derivation that the gradient of $\phi(\mathbf{R})$ is

$$\frac{\partial \phi}{\partial \mathbf{R}} = \mathbf{S}_{ur}^T \boldsymbol{\varepsilon},\ \boldsymbol{\varepsilon} = [\frac{1}{\widetilde{u}_1 - \widetilde{\mathbf{s}}_1^T \mathbf{R}} \quad \cdots \quad \frac{1}{\widetilde{u}_{N-n} - \widetilde{\mathbf{s}}_{N-n}^T \mathbf{R}}]^T, \tag{18}$$

and $\boldsymbol{\varepsilon}$ can be evaluated experimentally at the current point **R**. But this gradient can be obtained by conducting a special gradient experiment which is very similar to the one used for the gradient estimation of the original o.f. That is also the reason to split the new o.f. constraints in two ones (and hence the barrier function in two components) as we need an extra experiment for the rest of $N-n$ constraints in order to take advantage of the dimensionality of the map $\mathbf{S}_{ur}$. Therefore, the new gradient of the o.f. in (17) is

$$\begin{aligned}
&\left.\frac{\partial \widetilde{J}}{\partial \mathbf{R}}\right|_{\mathbf{R}=\mathbf{R}_j} = 2\mathbf{T}^T \mathbf{E}_j + \kappa\,(\mathbf{S}_{ur}^T \boldsymbol{\varepsilon}_G^1(\mathbf{R}_j) + \mathbf{S}_{ur}^T \boldsymbol{\varepsilon}_G^2(\mathbf{R}_j)),\\
&\boldsymbol{\varepsilon}_G^1(\mathbf{R}) = [\frac{1}{\widetilde{u}_1 - \widetilde{\mathbf{s}}_1^T \mathbf{R}} \quad \cdots \quad \frac{1}{\widetilde{u}_{N-n} - \widetilde{\mathbf{s}}_{N-n}^T \mathbf{R}}]^T,\\
&\boldsymbol{\varepsilon}_G^2(\mathbf{R}) = [\frac{1}{\widetilde{u}_{N-n+1} - \widetilde{\mathbf{s}}_{N-n+1}^T \mathbf{R}} \quad \cdots \quad \frac{1}{\widetilde{u}_{2(N-n)} - \widetilde{\mathbf{s}}_{2(N-n)}^T \mathbf{R}}]^T,
\end{aligned} \tag{19}$$

where $\boldsymbol{\varepsilon}_G^1(\mathbf{R}_j)$ holds for the first $N-n$ constraints, $\boldsymbol{\varepsilon}_G^2(\mathbf{R}_j)$ holds for the rest of $N-n$ constraints, and the subscript G stands for gradient. The gradient can be obtained in three special experiments at each iteration to be shown as follows.

In order to accelerate the convergence, the following estimate of the Hessian of the o.f. can be used for a Newton search:

$$\begin{aligned}
&\mathbf{H}(\mathbf{R}_j) = \left.\frac{\partial^2 \widetilde{J}}{\partial \mathbf{R}^2}\right|_{\mathbf{R}=\mathbf{R}_j} = 2\mathbf{T}^T \mathbf{T} + \kappa\,\mathbf{S}_{ur}^T \mathbf{S}_{ur}\,(\boldsymbol{\varepsilon}_H^1(\mathbf{R}_j) + \boldsymbol{\varepsilon}_H^2(\mathbf{R}_j)),\\
&\boldsymbol{\varepsilon}_H^1(\mathbf{R}) = [\frac{1}{(\widetilde{u}_1 - \widetilde{\mathbf{s}}_1^T \mathbf{R})^2} \quad \cdots \quad \frac{1}{(\widetilde{u}_{N-n} - \widetilde{\mathbf{s}}_{N-n}^T \mathbf{R})^2}]^T,\\
&\boldsymbol{\varepsilon}_H^2(\mathbf{R}) = [\frac{1}{(\widetilde{u}_{N-n+1} - \widetilde{\mathbf{s}}_{N-n+1}^T \mathbf{R})^2} \quad \cdots \quad \frac{1}{(\widetilde{u}_{2(N-n)} - \widetilde{\mathbf{s}}_{2(N-n)}^T \mathbf{R})^2}]^T,
\end{aligned} \tag{20}$$

where $\boldsymbol{\varepsilon}_H^1(\mathbf{R})$ and $\boldsymbol{\varepsilon}_H^2(\mathbf{R})$ are expressed for each of the two different sets of $N-n$ constraints. **T** and $\mathbf{S}_{ur}$ can be recorded from experiments, and both $\boldsymbol{\varepsilon}_H^1(\mathbf{R})$ and $\boldsymbol{\varepsilon}_H^2(\mathbf{R})$ can be evaluated after a normal experiment at the current **R** with no other extra effort. Here the subscript H stands for Hessian. A Gauss-Newton approximation of the Hessian can also be used.

The typical behavior of the IPB algorithm is a multiple stage optimization starting with an initial κ^0 and decreasing the current κ by a constant factor μ at each outer loop iteration (referred to as stage) until the desired accuracy is reached. At each stage, an iterative procedure consisting of a Newton or steepest descent search is conducted up to a specified accuracy level. The inner search is carried out in terms of the iterative procedure called iteration.

Our DDRTTA employs the IPB algorithm combined with a Newton search. A stage of the algorithm consists of two steps and an iteration of the algorithm consists of four sub-steps. The DDRTTA is summarized in terms of steps S1 and S2:

Step S1. Choose $\kappa^0 > 0,\ \mu > 1$, and start with a feasible point for $\mathbf{R}_0$ (the initial guess of **R**). Choose the upper and lower bounds for the control signal and generate the desired reference trajectory vector $\mathbf{Y}^d$. Choose the tolerances for stopping the inner and the outer search procedures, tol_N and tol_{IPB}. Set the outer iteration index for κ to $j_\kappa = 0$ and for **R** to $j = 0$. Assume that **T** and $\mathbf{S}_{ur}$ are estimated from measurements, and carry out the sub-steps S-s1 to S-s4:

- *Sub-step S-s1*. Choose γ_0 and reset the inner iteration index to $i = 0$. At each iteration of the search algorithm carry out the following sub-steps:
- *Sub-step S-s2*. Conduct a normal experiment with the current $\mathbf{R}_j$. Evaluate the o.f. in (17) and the vector variables $\boldsymbol{\varepsilon}_G^1, \boldsymbol{\varepsilon}_G^2, \boldsymbol{\varepsilon}_H^1, \boldsymbol{\varepsilon}_H^2$. Conduct a gradient experiment according to the approach given in Section III to find the gradient of the first three terms of the o.f.

in (17) which corresponds to $\mathbf{E}_j^T\mathbf{E}_j$. Conduct a gradient experiment in terms of Section III and apply (18) with $\boldsymbol{\varepsilon}=\boldsymbol{\varepsilon}_G^1$ to find the gradient which corresponds to half of the constraints. Conduct another gradient experiment as shown in Section III and apply (18) with $\boldsymbol{\varepsilon}=\boldsymbol{\varepsilon}_G^2$ to find the gradient which corresponds to the other half of the constraints.

- *Sub-step S-s3*. Compute the gradient using (19), evaluate the Hessian using (20) for $\mathbf{T}$, $\mathbf{S}_{ur}$ and $\boldsymbol{\varepsilon}_H^1,\boldsymbol{\varepsilon}_H^2$. Calculate the next reference input sequence using the following modified expression of (7):

$$\mathbf{R}_{j+1}=\mathbf{R}_j-\gamma_i(\mathbf{H}(\mathbf{R}_j))^{-1}\left.\frac{\partial\widetilde{J}}{\partial\mathbf{R}}\right|_{\mathbf{R}=\mathbf{R}_j}. \tag{21}$$

Set $i=i+1$, $j=j+1$.

- *Sub-step S-s4*. If the Newton search has converged in terms of a minimum tolerated decrease in the o.f. $J(\mathbf{R}_j)-J(\mathbf{R}_{j-1})<tol_N$, jump to next step, otherwise go to sub-step S-s2.

Step S2. If the search is satisfactory with the current κ^{j_κ}, $2(N-n)\kappa^{j_\kappa}<tol_{IPB}$ then stop. Otherwise set $\kappa^{j_k+1}=\kappa^{j_k}/\mu$, $j_\kappa=j_\kappa+1$, and jump to step S1.

The parameters tol_N, $tol_N>0$, and tol_{IPB}, $tol_{IPB}>0$, indicate the accepted tolerance for the increment of J and for the stop condition, respectively.

The simplest initial feasible point for $\mathbf{R}_0$ is to choose it equal to zero assuming zero initial conditions for the process. This approach can also be made suitable for nonzero initial conditions by choosing a corresponding $\mathbf{R}_0$. Therefore the constraints are not violated and the feasibility is preserved.

In the stochastic case, the constraints may be violated in the IPB algorithm due to the noise affecting the closed-loop CS. This in turn will drive the o.f. to infinity even if the deterministic counterpart of the signals (the control signal here) does not cross the feasible region boundaries. An alternate solution to this case is given in [51], with the following expression of the o.f. with penalty:

$$\begin{aligned}\widetilde{J}_p(\mathbf{R},\boldsymbol{\lambda})&=J(\mathbf{R})\\&+\frac{1}{2p}\sum_{h=1}^{c}\{[\max\{0,\lambda_h-p(\widetilde{u}_h-\widetilde{\mathbf{s}}_h^T\mathbf{R})\}]^2-\lambda_h^{\,2}\},\end{aligned} \tag{22}$$

where h is the constraint index, $h=1...c$, p, $p>0$, is a penalty parameter, and $\boldsymbol{\lambda}=[\lambda_1\ \ ...\ \ \lambda_c]^T\in\Re^c$ can be regarded as an estimate of the Lagrange multiplier vector. If $\{p_n\}_{n\geq 0}$ is a positive and strictly increasing sequence of penalty parameters and $\{\boldsymbol{\lambda}_n\}_{n\geq 0}$ is a bounded non-negative sequence of Lagrange multiplier vectors, then the minimum of $\{\widetilde{J}_{p_n}(\mathbf{R},\boldsymbol{\lambda}_n)\}_{n\geq 0}$ will converge to the solution to the constrained optimization problem (16). The minimization of (22) is carried out using a stochastic approximation algorithm and the gradient of the o.f. requires knowledge of the same gradient information as in the case of the IPB algorithm.

V. Case Study and Discussion of Results

The case study deals with the angular position of the experimental setup built around an INTECO DC servo system laboratory equipment [20]. The nonlinear state-space model of the process is

$$\begin{aligned}&m(t)=\begin{cases}k_{u,m}(u_a-u_c), & \text{if } u(t)\leq -u_c,\\ k_{u,m}[u(t)+u_a], & \text{if } -u_c<u(t)<-u_a,\\ 0, & \text{if } |u(t)|\leq u_a,\\ k_{u,m}[u(t)-u_a], & \text{if } u_a<u(t)<u_b,\\ k_{u,m}(u_b-u_a), & \text{if } u(t)\geq u_b,\end{cases}\\&\dot{\mathbf{x}}_P(t)=\begin{bmatrix}0 & 1\\ 0 & -1/T_\Sigma\end{bmatrix}\mathbf{x}_P(t)+\begin{bmatrix}0\\ k_{P1}/T_\Sigma\end{bmatrix}m(t)+\begin{bmatrix}1\\0\end{bmatrix}d(t),\\&y(t)=[1\quad 0]\,\mathbf{x}_P(t),\end{aligned} \tag{23}$$

where $t\in\Re$, $t\geq 0$ is the continuous time argument, u is a pulse width modulated duty cycle, d is the disturbance input, m is the output of the saturation and dead zone static nonlinearity (the first equation in (23)) with the parameters $k_{u,m}>0$ and $0<u_a<u_b,u_c$, and the parameters of the linear dynamics in (23) are k_{P1} and T_Σ. $\mathbf{x}_P(t)=[\alpha(t)\quad\omega(t)]^T$ is the state vector, $y(t)=\alpha(t)$ is the angular position and $\omega(t)$ is the angular speed.

The dynamic counterpart of the model has not been used in the tuning procedure. Only the dead zone is compensated using an inverse nonlinearity. Other effects such as viscous friction introduce the randomness in all measurements.

The values of parameters in (23) are: $u_a=0.15$, $u_b=u_c=1$, $k_{u,m}=1/(u_b-u_a)=1.17$ and $T_\Sigma=0.9\,\text{s}$. The process transfer function $P(s)=140/[s(1+0.9s)]$ is a good approximation of (23) according to [20]. $P(s)$ is used in tuning the continuous-time controller in terms of the Extended Symmetrical Optimum method [52] which uses a single design parameter in tuning equations for PI and PID controllers dedicated to many industrial servo system applications [53]–[63]. The discrete-time feedback PI controller results from the continuous-time one $C(s)=0.003(1+4s)/s$. The reference trajectory is prescribed in terms of the unit step response of a second-order normalized reference model with the transfer function $M(s)=\omega_n^2/(s^2+2\zeta\omega_n s+\omega_n^2)$, with $\omega_n=1$ rad/s and $\zeta=0.7$. An operational constraint is considered by imposing the control signal to be bounded within $-0.1\leq u\leq 0.2$. The sampling period is 0.1 s and the length of experiments is of $N=150$ samples. The relative degree of $T(q^{-1})$ is $n=1$ and the relative degree of $S_{ur}(q^{-1})$ is $m=0$.

The DDRTTA is applied in a stochastic framework for three IPB iterations corresponding to $\kappa^0=0.5$, $\kappa^1=0.5$ and $\kappa^2=0.5$. At each IPB iteration the corresponding initial step-scaling coefficient was chosen as $\gamma_0=20$, $\gamma_0=5$ and $\gamma_0=2$. The stochastic setup was created by injecting zero mean white noise as an output disturbance (see Fig. 1), with variance $\sigma_v^2=0.04$. The DDRTTA was run for 50 times. The initial reference input versus time is illustrated in Fig. 2. The initial position y and control signal u versus time are also given in Fig. 2 along with their final (after optimization) responses. The solution based on (22) was not used at this point. Instead a tradeoff was made to decrease κ at a slow rate and not too much in order to prevent the violation of the constraints. Even in this case, the final results show that the final output response is close to the reference model trajectory while the constraints on u are fulfilled. So a nearly optimal solution is attained showing the effectiveness of DDRTTA.

The residuals of the objective function are defined as $\bar{\tilde{J}}_i^j=\tilde{J}_i^j-avg(\tilde{J}_i)$, where $1\le i\le 14$ is the iteration index, $1\le j\le 50$ is the algorithm run index, and $avg(\tilde{J}_i)$ is the sample average over all the 50 runs at the i^{th} iteration. The residuals are shown in Fig. 3.

Our algorithm optimizes the reference input vector for 149 variables and 298 constraints. It requires four experiments per iteration (one normal one and three gradient ones) for 14 iterations.

The decrease in the o.f. is substantial after only a small number of iterations that could be accepted as a reasonable solution without experimenting further on the CS.

The experimental validation of the proposed DDRTTA is next presented. The controller, reference trajectory and initial reference input are considered as in the simulations. Two situations are illustrated here corresponding to an unconstrained optimization and to a constrained one with the control signal constraints as in the simulations. The IPB algorithm is used in a simplified steepest-descent fashion in the stochastic setting because the randomness is inherent in the real-world application. The same approach is used in the unconstrained case. The same initial feasible solution was chosen as a starting point in both optimization algorithms.

The IPB was applied for three iterations corresponding to $\kappa^0=0.5$, $\kappa^1=0.05$ and $\kappa^2=0.005$. At each iteration the corresponding step-scaling parameters of DDRTTA were chosen according to (13), with $\gamma_0=20$, $\gamma_0=10$ and $\gamma_0=5$, respectively. The step-scaling parameter $\gamma_0=20$ was set in the unconstrained case. In both cases the unity matrix was used for $\mathbf{H}(\mathbf{R}_j)$. The compared results before and after the applications of our DDRTTA are shown in Figs. 4, 5 and 6. The same approach as in the simulation case study was used for a smooth decrease of κ to get a nearly optimal solution.

The improvements are visible even after a small number of iterations, both for the digital simulation results and the experimental ones. Even if the minimum is not reached the decrease of the o.f. is remarkable.

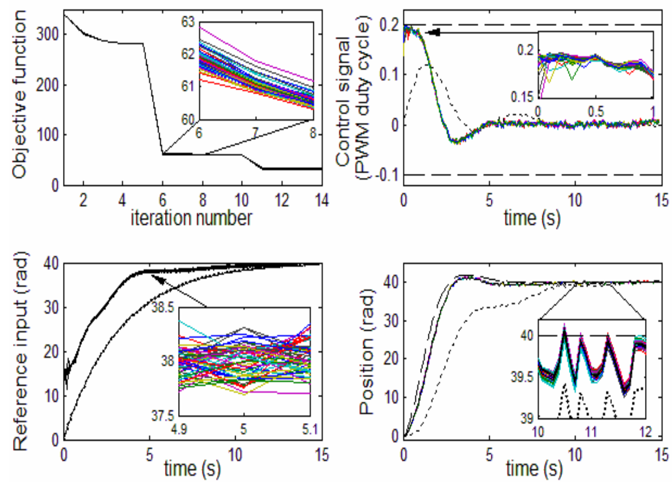

Fig. 2. Simulation results for 50 runs of DDRTTA, upper-left: the objective function, upper right: the initial control signal (dotted), the upper and lower bounds (thick dashed grey), lower left: the initial reference input (dotted) and the final reference inputs (solid), lower right: initial position response (dotted), reference model (dashed), final responses (solid)

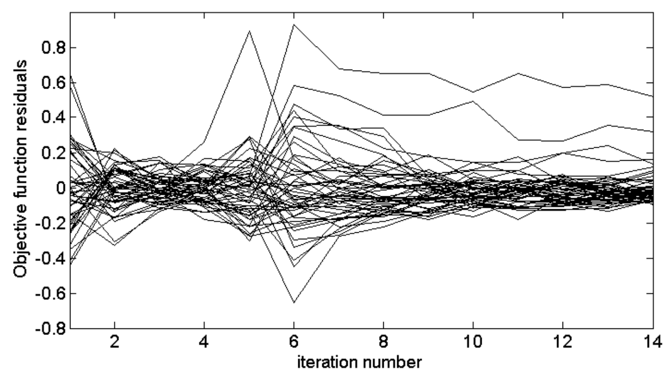

Fig. 3. Objective function residuals across the iterations for 50 runs of the DDRTTA algorithm.

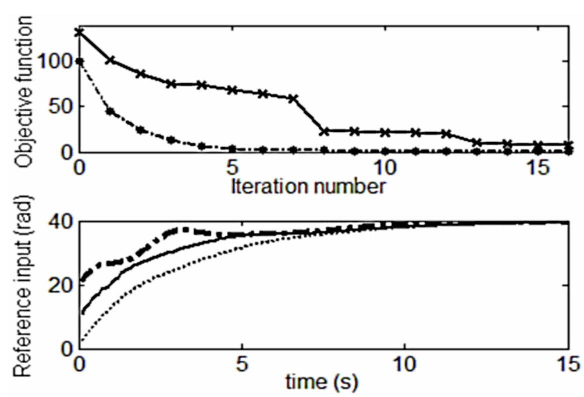

Fig. 4. Experimental results expressed as o.f. evolution: unconstrained case (dash-dot), constrained (solid), and as optimized reference input: initial (dotted), unconstrained case (dash-dot), constrained case (solid).

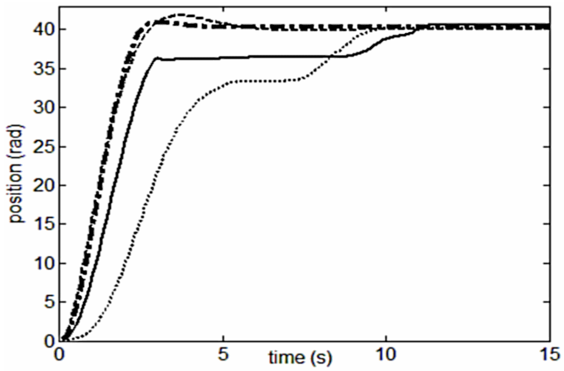

Fig. 5. Experimental results expressed as position response: initial (dotted), unconstrained case (dash-dot), constrained case (solid) and reference trajectory (dash).

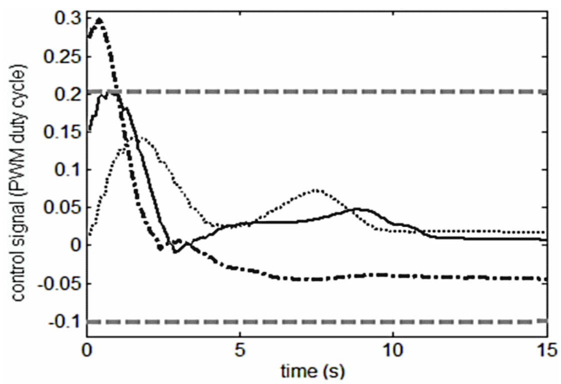

Fig. 6. Experimental results expressed as control signal responses: initial (dotted), unconstrained case (dash-dot), constrained case (solid), upper and lower bounds (thick dashed grey).

VI. Conclusion

This paper has proposed a data-driven algorithm which solves an optimal control problem focused on reference trajectory tracking. The search procedure starts with a feasible solution and it is entirely model-free data-driven in the steepest descent case.

One of the advantages of the new DDRTTA is that it works for smooth nonlinear systems around some operating points. Also, for time-varying systems for which the variation time scale is larger than the algorithm's time scale, otherwise it may not converge to a solution.

Another advantage of the proposed algorithm is that the closed-loop CS is not affected while solving the trajectory tracking problem. Starting with a given closed-loop controller (which can be obtained in terms of a model-based or model-free design) the stability and the disturbance rejection are preserved. The latter can also be regarded as a shortcoming when the disturbance characteristics may change therefore requiring controller retuning. The change in the controller parameters specific to other tuning techniques including IFT which is important in order to achieve the bumpless transfer between controllers is mitigated by our approach.

One limitation of the DDRTTA may be caused by the abnormal experimenting regimes or by the length of the optimization procedure or even by the duration of the experiments for the slow processes. However, assuming that the closed-loop CS is stable, the safety of the experiments is guaranteed around the operating point.

The quality of the gradient estimate influences the length of the search procedure. Finite differences approximation estimation of the gradient would be a poor choice in the case of noisy measurements. Therefore the compromise to the small number of experiments per iteration and to the quality of the gradient information would require the use of an identified model for the closed-loop system. This approach eliminates three gradient experiments with the obvious advantage of trade-off to number of gradient experiments and gradient accuracy. In addition, the problem of the abnormal operating regimes is solved. The accuracy of the estimated gradient is expected to be good in case of smooth nonlinear processes, and it is a future research topic.

Our algorithm can be generalized by considering other data-driven optimization approaches to controller design combined with ILC to optimize the reference input. The main advantage of such combinations is that they can lead to automated tools for controller design with benefits in complex industrial systems applications.

References

[1] M. Ikeda, Y. Fujisaki, and N. Hayashi, "A model-less algorithm for tracking control based on input-output data," *Nonlinear Anal. Theory Methods Appl.*, vol. 47, pp. 1953–1960, Aug. 2001.

[2] I. Markovsky and P. Rapisarda, "Data-driven simulation and control," *Int. J. Control*, vol. 81, pp. 1946–1959, Dec. 2008.

[3] M. Helle and H. Saxén, "Data-driven analysis of sulfur flows and behavior in the blast furnace," *Steel Res. Int.*, vol. 79, pp. 671–677, Sep. 2008.

[4] J. Zeng, C. Gao, and H. Su, "Data-driven predictive control for blast furnace ironmaking process," *Comput. Chem. Eng.*, vol. 34, pp. 1854–1862, Nov. 2010.

[5] D. Wang, "Robust data-driven modeling approach for real-time final product quality prediction in batch process operation," *IEEE Trans. Ind. Informat.*, vol. 7, pp. 371–377, May 2011.

[6] C. Gao, L. Jian, X. Liu, and J. Chen, "Data-driven modeling based on Volterra series for multidimensional blast furnace system," *IEEE Trans. Neural Netw.*, vol. 22, pp. 2272–2283, Dec. 2011.

[7] H. Hjalmarsson, M. Gevers, S. Gunnarsson, and O. Lequin, "Iterative feedback tuning: theory and applications," *IEEE Control Syst. Mag.*, vol. 18, pp. 26–41, Aug. 1998.

[8] A Karimi, L Miskovic, and D. Bonvin, "Iterative correlation-based controller tuning," *Int. J. Adapt. Control Signal Process.*, vol. 18, pp: 645–664, Oct. 2004.

[9] L. C. Kammer, "Stability assessment for cautious iterative controller tuning," *Automatica*, vol. 41, pp. 1829–1834, Oct. 2005.

[10] R.-E. Precup, C. Borchescu, M.-B. Rădac, S. Preitl, C.-A. Dragoş, E. M. Petriu, and J. K. Tar, "Implementation and signal processing aspects of iterative regression tuning," in *Proc. 2010 IEEE International Symposium on Industrial Electronics (ISIE 2010)*, Bari, Italy, 2010, pp. 1657–1662.

[11] 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.

[12] M.-B. Rădac, R.-E. Precup, E. M. Petriu, and S. Preitl, "Application of IFT and SPSA to servo system control," *IEEE Trans. Neural Netw.*, vol. 22, pp. 2363–2375, Dec. 2011.

[13] J. K. Bennighof, S.-H. Chang, and M. Subramaniam, "Minimum time pulse response based control of flexible structure," *J. Guid. Control Dyn.*, vol. 16, pp. 874–881, Oct. 1993.

[14] G. Shi and R. E. Skelton, "Markov data-based LQG control," *J. Dyn. Syst. Meas. Control*, vol. 122, pp. 551–559, Sep. 2000.

[15] R. Kadali, B. Huang, and A. Rossiter, "A data driven subspace approach to predictive controller design," *Control Eng. Prect.*, vol. 11, pp. 261–278, Mar. 2003.

[16] X. Wang, B. Huang, and T. Chen, "Data-driven predictive control for solid oxide fuel cells," *J. Process Control*, vol. 17, pp. 103–114, Feb. 2007.

[17] X. Lu, H. Chen, P. Wang, and B. Gao, "Design of a data-driven predictive controller for start-up process of AMT vehicles," *IEEE Trans. Neural Netw.*, vol. 22, pp. 2201–2212, Dec. 2011.

[18] W. Favoreel, B. De Moor, P. van Overschee, and M. Gevers, "Model-free subspace-based LQG-design," in *Proc. 1999 American Control Conference*, San Diego, CA, USA, 1999, vol. 5, pp. 3372–3376.

[19] M. C. Campi, A. Lecchini, and S. M. Savaresi, "Virtual reference feedback tuning: a direct method for the design of feedback controllers," *Automatica*, vol. 38, pp. 1337–1346, Aug. 2002.

[20] M.-B. Rădac, R.-E. Precup, E. M. Petriu, S. Preitl, and R.-C. David, "Stable iterative feedback tuning method for servo systems," in *Proc. 20th IEEE International Symposium on Industrial Electronics (ISIE 2011)*, Gdansk, Poland, 2011, pp. 1943–1948.

[21] D. A. Bristow, M. Tharayil, and A. G. Alleyne, "A survey of iterative learning control," *IEEE Control Syst. Mag.*, vol. 26, pp. 96–114, Jun. 2006.

[22] H.-S. Ahn, Y. Chen, and K. L. Moore, "Iterative learning control: brief survey and categorization," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 37, pp. 1109–1121, Nov. 2007.

[23] D. H. Owens and J. Hätönen, "Iterative learning control – an optimization paradigm," *Annu. Rev. Control*, vol. 29, pp. 57–70, Apr. 2005.

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

[25] M. Norrlöf and S. Gunnarsson, "Time and frequency domain convergence properties in iterative learning control," *Int. J. Control*, vol. 75, pp. 1114–1126, Sep. 2002.

[26] 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.

[27] H.-F. Chen and H.-T. Fang, "Output tracking for nonlinear stochastic systems by iterative learning control," *IEEE Trans. Autom. Control*, vol. 49, pp. 583–588, Apr. 2004.

[28] P. Baranyi and Y. Yam, "Singular value-based approximation with Takagi-Sugeno type fuzzy rule base," in *Proc. 6th IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'97)*, Barcelona, Spain, 1997, vol. 1, pp. 265–270.

[29] T. Orlowska-Kowalska, K. Szabat, and K. Jaszczak, "The influence of parameters and structure of PI-type fuzzy-logic controller on DC drive system dynamics," *Fuzzy Sets Syst.*, vol. 131, pp. 251–264, Oct. 2002.

[30] Z. Petres, P. Baranyi, P. Korondi, and H. Hashimoto, "Trajectory tracking by TP model transformation: Case study of a benchmark problem," *IEEE Trans. Ind. Electron.*, vol. 54, pp. 1654–1663, Jun. 2007.

[31] S. Blažič and I. Škrjanc, "Design and stability analysis of fuzzy model-based predictive control - a case study," *J. Intell. Robotic Syst.*, vol. 49, pp. 279–292, Jul. 2007.

[32] S. Blažič, I. Škrjanc, S. Gerkšič, G. Dolanc, S. Strmčnik, M. B. Hadjiski, and A. Stathaki, "Online fuzzy identification for an intelligent controller based on a simple platform," *Eng. Appl. Artif. Intell.*, vol. 22, pp. 628–638, Jun. 2009.

[33] J. Vaščák and L. Madarász, "Adaptation of fuzzy cognitive maps – a comparison study," *Acta Polyt. Hung.*, vol. 7, pp. 109–122, Sep. 2010.

[34] Z. C. Johanyák, "Student evaluation based on fuzzy rule interpolation," *Int. J. Artif. Intell.*, vol. 5, pp. 37–55, Sep. 2010.

[35] Z. C. Johanyák, "Survey on five fuzzy inference-based student evaluation methods," in *Computational Intelligence in Engineering*, I. J. Rudass, J. Fodor, and J. Kacprzyk, Eds., Berlin, Heidelberg: Springer-Verlag, Studies in Computational Intelligence, vol. 313, 2010, pp. 219–228.

[36] O. Linda and M. Manic, "Uncertainty-robust design of interval type-2 fuzzy logic controller for delta parallel robot," *IEEE Trans. Ind. Informat.*, vol. 7, pp. 661–670, Nov. 2011.

[37] H. Yu, T. Xie, S. Paszczynski, and B. M. Wilamowski, "Advantages of radial basis function networks for dynamic system design," *IEEE Trans. Ind. Electron.*, vol. 58, pp. 5438–5450, Dec 2011.

[38] X. Zhang, S. Hu, D. Chen, and X. Li, "Fast covariance matching with fuzzy genetic algorithm," *IEEE Trans. Ind. Informat.*, vol. 8, pp. 148–157, Feb. 2012.

[39] C. Gao, L. Jian, and S. Luo, "Modeling of the thermal state change of blast furnace hearth with support vector machines," *IEEE Trans. Ind. Electron.*, vol. 59, pp. 1134–1145, Feb. 2012.

[40] L. Rutkowski, A. Przybyl, and K. Cpalka, "Novel online speed profile generation for industrial machine tool based on flexible neuro-fuzzy approximation," *IEEE Trans. Ind. Electron.*, vol. 59, pp. 1238–1247, Feb. 2012.

[41] Z. Gao and S. X. Ding, "Actuator fault robust estimation and fault-tolerant control for a class of nonlinear descriptor systems," *Automatica*, vol. 43, pp. 912–920, May 2007.

[42] F. Pettersson, H. Saxén, and K. Deb, "Genetic-algorithm based multi-criteria optimization of ironmaking in the blast furnace," *Mater. Manuf. Process.*, vol. 24, pp. 343–349, Mar. 2009.

[43] Na. Pham, A. Malinowski, and T. Bartczak, "Comparative study of derivative free optimization algorithms," *IEEE Trans. Ind. Informat.*, pp. 592–600, Nov. 2011.

[44] H. Liu and S. Li, "Speed control for PMSM servo system using predictive functional control and extended state observer," *IEEE Trans. Ind. Electron.*, vol. 59, pp. 1171–1183, Feb. 2012.

[45] R.-E. Precup, R.-C. David, E. M. Petriu, S. Preitl, and M.-B. Rădac, "Fuzzy control systems with reduced parametric sensitivity based on simulated annealing," *IEEE Trans. Ind. Electron.*, vol. 59, pp. 3049–3061, Aug. 2012.

[46] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT Press: Cambridge, MA, 1998.

[47] B. Scherer, "Asynchronous neuron computing for optimal control and reinforcement learning with large states paces," *Neurocomputing*, vol. 63, pp. 229–251, Jan. 2005.

[48] R. C. Hsu, C.-T. Liu, and D.-Y. Chan, "A reinforcement-learning-based assisted power management with QoR provisioning for human-electric hybrid bicycle," *IEEE Trans. Ind. Electron.*, vol. 59, pp. 3350–3359, Aug. 2012.

[49] 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.

[50] S. Mishra, U. Topcu, and M. Tomizuka, "Optimization-based constrained iterative learning control," *IEEE Trans. Contr. Syst. Technol.*, vol. 19, pp. 1613–1621, Nov. 2011.

[51] 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.

[52] S. Preitl and R.-E. Precup, "An extension of tuning relations after symmetrical optimum method for PI and PID controllers," *Automatica*, vol. 35, pp. 1731–1736, Oct. 1999.

[53] F. Pettersson, N. Chakraborti, and H. Saxén, "A genetic algorithms based multi-objective neural net applied to noisy blast furnace data," *Appl. Soft Comput.*, vol. 7, pp. 387–397, Jan. 2007.

[54] T. Orlowska-Kowalska and K. Szabat, "Damping of torsional vibrations in two-mass system using adaptive sliding neuro-fuzzy approach," *IEEE Trans. Ind. Informat.*, vol. 4, pp. 47–57, Feb. 2008.

[55] Z. Gao, X. Dai, T. Breikin, and H. Wang, "Novel parameter identification by using a high-gain observer with application to a gas turbine engine," *IEEE Trans. Ind. Informat.*, vol. 4, pp. 271–279, Nov. 2008.

[56] F. Pettersson, H. Saxén, and K. Deb, "Genetic-algorithm based multi-criteria optimization of ironmaking in the blast furnace," *Materials and Manufacturing Processes*, vol. 24, pp. 343–349, Mar. 2009.

[57] J. A. Iglesias, P. Angelov, A. Ledezma, and A. Sanchis, "Evolving classification of agents' behaviors: a general approach," *Evolving Syst.*, vol. 1, pp. 161–171, Oct. 2010.

[58] S.-H. Hur, R. Katebi, and A. Taylor, "Modeling and control of a plastic film manufacturing web process," *IEEE Trans. Ind. Informat.*, vol. 7, pp. 171–178, May 2011.
[59] S. Sakaino, T. Sato, and K. Ohnishi, "Multi-DOF micro-macro bilateral controller using oblique coordinate control *IEEE Trans. Ind. Informat.*, vol. 7, pp. 446–454, Aug. 2011.
[60] R. Liu, M. Niu, L. Tang, and L. Jiao, "Adaptive particle swarm optimization based artificial immune network classification algorithm," *Int. J. Artif. Intell.*, vol. 7, pp. 151–164, Oct. 2011.
[61] F.-J. Lin, P.-H. Chou, C.-S. Chen, and Y.-S. Lin, "DSP-based cross-coupled synchronous control for dual linear motors via intelligent complementary sliding mode control," *IEEE Trans. Ind. Electron.*, vol. 59, pp. 1061–1073, Feb. 2012.
[62] J. Vaščák and M. Paľa, "Adaptation of fuzzy cognitive maps for navigation purposes by migration algorithms," *Int. J. Artif. Intell.*, vol. 8, pp. 20–37, Mar. 2012.
[63] P. A. Sente, F. M. Labrique, and P. J. Alexandre, "Efficient control of a piezoelectric linear actuator embedded into a servo-valve for aeronautic applications," *IEEE Trans. Ind. Electron.*, vol. 59, pp. 1971–1979, Apr. 2012.

Mircea-Bogdan Rădac (M'12) received the Dipl.Ing. degree in systems and computer engineering and the Ph.D. degree in systems engineering from the "Politehnica" University of Timisoara (PUT), Timisoara, Romania, in 2008 and 2011, respectively. He has been working towards the Doctoral degree at PUT since 2008.

He is currently with PUT, where he became an Assistant Lecturer with the Department of Automation and Applied Informatics in 2012. He is the co-author of more than 20 papers published in scientific journals, refereed conference proceedings, and contributions to books. His current research interests include control structures and algorithms with focus on iterative methods in control design and optimization.

Dr. Rădac is a member of the Romanian Society of Control Engineering and Technical Informatics.

Radu-Emil Precup (M'03–SM'07) received the Dipl.Ing. (with honors) degree in automation and computers from the "Traian Vuia" Polytechnic Institute of Timisoara, Timisoara, Romania, the Diploma degree in mathematics from the West University of Timisoara, Timisoara, and the Ph.D. degree in automatic systems from the "Politehnica" University of Timisoara (PUT), Timisoara, in 1987, 1993, and 1996, respectively.

He is currently with PUT, where he became a Professor with the Department of Automation and Applied Informatics in 2000. He is an Honorary Professor with the Óbuda University, Budapest, Hungary. He is the author or coauthor of more than 150 papers published in scientific journals, refereed conference proceedings, and contributions to books. His current research interests include intelligent control systems, data-based control, and nature-inspired algorithms for optimization.

Prof. Precup is a member of the Subcommittee on Computational Intelligence as part of the IEEE Industrial Electronics Society.

Emil M. Petriu (M'86–SM'88–F'01) received Dipl.Eng. and Dr.Eng. degrees in Electrical Engineering from the Polytechnic Institute of Timisoara, Timisoara, Romania.

He is a is a University Research Chair Professor with the School of Electrical Engineering and Computer Science, at the University of Ottawa, Ottawa, ON, Canada. His research interests include multisensor systems, soft computing, biology-inspired robot sensing, and human-computer symbiosis.

Prof. Petriu is a Fellow of the Canadian Academy of Engineering and the Engineering Institute of Canada. He was a co-recipient of the 2003 IEEE Donald G. Fink Prize Paper Award.

Stefan Preitl (M'03–SM'07) received the Dipl.Ing. (with honors) degree in electrical engineering and the Ph.D. degree in measurement techniques from the "Traian Vuia" Polytechnic Institute of Timisoara, Timisoara, Romania, in 1966 and 1983, respectively.

He is currently with the "Politehnica" University of Timisoara, Timisoara, where he became a Professor in the Department of Automation and Applied Informatics in 1992. He is an Honorary Professor with the Óbuda University, Budapest, Hungary. He is the author or co-author of more than 200 papers published in various scientific journals, refereed conference proceedings, and books in the field of automatic control. His current research interests include conventional and advanced structures and algorithms for automatic control applied to power or servo systems, control systems of electrical drives, methodical aspects of teaching, and development of computer-assisted education.

Prof. Preitl is a member of the International Federation of Automatic Control Technical Committee on Control Design.

Claudia-Adina Dragoş received the Dipl.Ing. degree in systems and computer engineering, the M.Sc. degree in automatic systems, and the Ph.D. degree in systems engineering from the "Politehnica" University of Timisoara (PUT), Timisoara, Romania, in 2007, 2009 and 2011, respectively. She has been working towards the Doctoral degree at PUT since 2008.

She is currently with PUT, where he became an Assistant Lecturer with the Department of Automation and Applied Informatics in 2012. She is the co-author of more than 30 papers published in scientific journals, refereed conference proceedings, and contributions to books. Her current research interests include control structures and algorithms focused on fuzzy, predictive and adaptive control, and systems modeling and optimization.

Dr. Dragoş is a member of the Romanian Society of Control Engineering and Technical Informatics.