Rational basis functions in iterative learning control - With experimental verification on a motion system

Rational Basis Functions in Iterative Learning Control—With Experimental Verification on a Motion System

Joost Bolder and Tom Oomen

***Abstract*—Iterative learning control (ILC) approaches often exhibit poor extrapolation properties with respect to exogenous signals, such as setpoint variations. This brief introduces rational basis functions in ILC. Such rational basis functions have the potential to both increase performance and enhance the extrapolation properties. The key difficulty that is associated with these rational basis functions lies in a significantly more complex optimization problem when compared with using preexisting polynomial basis functions. In this brief, a new iterative optimization algorithm is proposed that enables the use of rational basis functions in ILC for single-input single-output systems. An experimental case study confirms the advantages of rational basis functions compared with preexisting results, as well as the effectiveness of the proposed iterative algorithm.**

***Index Terms*—Basis functions, iterative learning control (ILC), optimal control.**

I. Introduction

Learning control is used in many motion systems. Examples include additive manufacturing machines [1], [2], robotic arms [3], printing systems [4], pick and place machines, electron microscopes, and wafer stages [5]–[7]. The often repetitive tasks for these systems typically vary to some degree to address tolerances in the products being processed.

Iterative learning control (ILC) [8] can significantly enhance the performance of systems that perform repeated tasks. After each repetition, the command signal for the next repetition is updated by learning from past executions. A key assumption in ILC is that the task of the system is invariant under the repetitions. As a consequence, the learned command signal is optimal for the specific task only. In general, extrapolation of the learned command signal to other tasks leads to a significant performance deterioration [6].

Several approaches have been proposed to enhance the extrapolation properties of ILC to a class of reference signals. In [5], a segmented approach to ILC is presented and applied to a wafer stage. This approach is further extended in [2], where the complete task is divided into subtasks that are learned individually. The use of such a signal library is restricting in the sense that tasks are required to consist of standardized building blocks. Instead of using a signal library, in [9]–[11], the extrapolation properties of ILC are enhanced through the use of basis functions. These basis functions can be used to parameterize the ILC command signal in terms of the task.

The preexisting results [4], [6] employ so-called polynomial basis functions. These polynomial basis functions can be interpreted as parameterizing the command signal in terms of the reference using a finite impulse response (FIR) filter. Importantly, such polynomial basis functions retain the analytic solution of the ILC as obtained in [8]. In [6], the polynomial basis functions in [12] are implemented in the ILC framework of [9] and successfully applied to an industrial wafer stage system, whereas in [4] an application to a wide-format printer application is reported. Finally, extensions of the approach toward input shaping are presented in [13].

Although the use of polynomial basis functions enhances the extrapolation properties of ILC algorithms, the polynomial nature of the basis functions severely limits the achievable performance and extrapolation properties. The basis functions typically constitute an approximate model inverse of the true system [14], [15]. The use of polynomial basis functions implies that a perfect inverse can be obtained only if the system has a unit numerator. Since many physical systems are modeled using rational models, containing both poles and zeros, this implies that existing results necessarily lead to undermodeling in the model inverses. Consequently, both the achievable performance and extrapolation properties are limited.

This brief aims to introduce a new ILC framework that can achieve improved performance and extrapolation properties for the class of single-input single-output rational systems. To this end, this brief introduces rational basis functions in ILC. The key feature is that both the numerator and denominator of the rational structure are parameterized, hence allowing both zeros and poles in obtaining the model inverse. The technical difficulty associated with these basis functions is that the analytic solution of standard optimal ILC [8] and the basis function approach in [9] is lost. In fact, the resulting optimization problem is nonconvex, in general.

The main and novel contribution of this brief is the introduction of rational basis functions in ILC, and a new parameter update solution that resorts to a sequence of optimization problems with an analytic solution. Interestingly, the results in [6], [9], and [13] are directly recovered as a special case of the novel results. The proposed solution has strong connections to common algorithms in both time domain system identification [16] and frequency domain system identification [17], [18].

The notation that is used in this brief is introduced in Section II. In Section III, the problem formulation is formally stated. Then, in Section IV, the new parameterization is proposed, followed by an analysis of its consequences for the optimization problem (Section IV-A). Section IV-B contains a novel iterative solution to the optimization problem, which constitutes the main contribution of this brief. Section V establishes connections to preexisting results that employ basis functions in ILC. In Section VI, an experimental case study is presented that reveals the advantages of employing rational basis functions and efficacy of the proposed iterative solution.

II. Preliminaries

A discrete-time transfer function is denoted as $\boldsymbol{H}(z)$, with z a complex indeterminate. The ith element of a vector θ is expressed as $\theta[i]$. A matrix $B \in \mathbb{R}^{n\times n}$ is defined positive (semi)definite iff

Manuscript received February 28, 2014; accepted May 9, 2014. Date of publication June 17, 2014; date of current version February 11, 2015. Manuscript received in final form May 26, 2014. This work is supported by Océ Technologies, and by the Innovational Research Incentives Scheme under the VENI grant "Precision Motion: Beyond the Nanometer" (no. 13073) awarded by NWO (The Netherlands Organisation for Scientific Research) and STW (Dutch Science Foundation). Recommended by Associate Editor A. Behal.

The authors are with the Control Systems Technology Group, Department of Mechanical Engineering, Eindhoven University of Technology, Eindhoven 5600 MB, The Netherlands (e-mail: jbolder@gmail.com; t.a.e.oomen@tue.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCST.2014.2327578

$x^T Bx \geq 0, \forall x \neq 0 \in \mathbb{R}^n$ and is denoted as $B \succeq 0$. For a vector x, the weighted 2-norm is $||x||_W = x^T W x$.

All signals and systems are discrete time and often implicitly assumed of length n. Given a system $\boldsymbol{H}(z)$, and input and output vectors $u, y \in \mathbb{R}^{n\times 1}$. Let $h(t), t \in \mathbb{Z}$ be the infinite-time impulse response vector of $\boldsymbol{H}(z)$. Then, the finite-time response of $\boldsymbol{H}$ to u is given by the truncated convolution

$$y[t] = \sum_{l=1-n}^{t} h(l)u[t-l]$$

with $0 \leq t < n$, and zero initial conditions. This finite-time convolution is written as

$$\underbrace{\begin{bmatrix} y[0] \\ y[1] \\ \vdots \\ y[n-1] \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} h(0) & h(-1) & \dots & h(1-n) \\ h(1) & h(0) & \dots & h(2-n) \\ \vdots & & \ddots & \vdots \\ h(n-1) & h(n-2) & \dots & h(0) \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} u[0] \\ u[1] \\ \vdots \\ u[n-1] \end{bmatrix}}_{u}$$

with H the convolution matrix corresponding to $\boldsymbol{H}(z)$, and u, y the input and output vectors. Note that $\boldsymbol{H}(z)$ is not restricted to be a causal system.

Given a transfer function with parametric coefficients

$$\boldsymbol{H}(\theta, z) = \sum_{i=1}^{m} \boldsymbol{\xi_i}(z)\theta[i]$$

with parameters $\theta \in \mathbb{R}^{m\times 1}$, and basis functions $\boldsymbol{\xi_i}(z)$, here $i = 1, 2, \dots, m$. The finite-time response of $\boldsymbol{H}$ to input u is given by

$$y = \Psi_{Hu}\theta \tag{1}$$

with $\Psi_{Hu} = [\xi_1 u, \xi_2 u, \dots, \xi_m u] \in \mathbb{R}^{n\times m}$, here $\xi_i \in \mathbb{R}^{n\times n}$ are the convolution matrices corresponding to $\boldsymbol{\xi_i}(z)$. Note that (1) is equivalent to $y = H(\theta)u$, with $H(\theta)$ the convolution matrix of $\boldsymbol{H}(\theta, z)$.

As an example, let $n = 4$, $\boldsymbol{\xi_1}(z) = 1$ and $\boldsymbol{\xi_2}(z) = z^{-1}$, then $\boldsymbol{H}(\theta, z) = \theta[1] + z^{-1}\theta[2]$, and accordingly

$$H(\theta) = \begin{bmatrix} \theta[1] & 0 & 0 & 0 \\ \theta[2] & \theta[1] & 0 & 0 \\ 0 & \theta[2] & \theta[1] & 0 \\ 0 & 0 & \theta[2] & \theta[1] \end{bmatrix}, \quad \Psi_{Hu} = \begin{bmatrix} u(1) & 0 \\ u(2) & u(1) \\ u(3) & u(2) \\ u(4) & u(3) \end{bmatrix}.$$

III. Problem Formulation

In this section, the problem addressed in this brief is defined in detail. First, in Section III-A, the general ILC setup is introduced. Then, in Section III-B, optimization-based ILC is introduced. This is further tailored toward polynomial basis functions in Section III-C, followed by a definition of the problem that is addressed in this brief.

A. Problem Setup

The considered ILC setup is shown in Fig. 1. The setup consists of a feedback controller C, and system P_0. Both are assumed linear time invariant (LTI), causal, and single-input single-output. During an experiment with index j and length n, the reference r and system output y_j are measured. The feedforward signal is denoted f_j. Note from Fig. 1 that

$$e_j = S_0 r - P_0 S_0 f_j \tag{2}$$

with $S_0 := (I + CP_0)^{-1}$ the sensitivity. In ILC, the feedforward is generated by learning from measured data of previous experiments, also called trials. The objective is to minimize e_{j+1}, i.e., the predicted tracking error for the next experiment. From (2), it follows that:

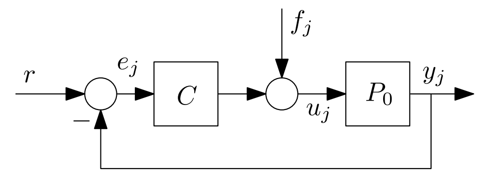

Fig. 1. ILC setup.

$$e_{j+1} = S_0 r - P_0 S_0 f_{j+1}. \tag{3}$$

Since r is constant, $S_0 r$ is eliminated from (2) and (3), yielding the error propagation from trial j to trial $j+1$

$$e_{j+1} = e_j - P_0 S_0 (f_{j+1} - f_j). \tag{4}$$

B. Norm-Optimal ILC

Norm-optimal ILC is an important class of ILC algorithms, where f_{j+1} is determined from the solution of an optimization problem, see [8], [19], [20]. Further extension, namely constrained optimization is considered in [21] and [22], which can for instance be used to prevent actuator saturation.

The optimization criterion in norm-optimal ILC is typically defined as follows.

Definition 1 (Norm-Optimal ILC): The optimization criterion for norm-optimal ILC algorithms is given by

$$\mathcal{J}(f_{j+1}) := ||e_{j+1}||_{W_e} + ||f_{j+1}||_{W_f} + ||f_{j+1} - f_j||_{W_{\Delta f}} \tag{5}$$

with $W_e \succ 0$, and $W_f, W_{\Delta f} \succeq 0$.

In (5), $W_e \succ 0$, and $W_f, W_{\Delta f} \succeq 0$ are user-defined weighting matrices to specify performance and robustness objectives, including: 1) robustness with respect to model uncertainty (W_f) and 2) convergence speed and sensitivity to trial varying disturbances ($W_{\Delta f}$). The corresponding feedforward update is given by

$$f_{j+1} = \arg\min_{f_{j+1}} \mathcal{J}(f_{j+1}). \tag{6}$$

The solution to (6) can be computed analytically from measurements e_j and f_j, given a model PS, since (5) is a quadratic function in f_{j+1}. The advantage of using a model in ILC in comparison to model-based feedback approaches is the significant performance improvements enabled by noncausal filtering operations in the time-domain as explained in [23].

In view of (3), the norm-optimal ILC computes a command signal f_j that is optimal in (5) for a specific reference trajectory r. As a result, changing r implies that the command signal f_j is not optimal in general. To introduce extrapolation capabilities in ILC, basis functions are introduced in Section II-C.

C. Norm-Optimal ILC With Polynomial Basis Functions

In [9], basis functions have been introduced in ILC that are of the form

$$f_j = \Psi\theta_j \tag{7}$$

where f_j is a linear combination of user-selected vectors $\Psi = [\psi_1, \psi_2, \dots, \psi_m]$. Notice that the basis functions Ψ in (7) are in so-called lifted notation. The basis functions in (7) encompass standard norm-optimal ILC with $\Psi = I$. Only specific choices enhance the extrapolation properties. The essence of enhancing extrapolation of the ILC command signal to different tasks lies in choosing f_j to be a function of r. Therefore, let $f_j = F(\theta_j)r$. Subsequent substitution into (2) yields

$$e_j = S_0 r - P_0 S_0 F(\theta_j) r = (I - P_0 F(\theta_j)) S_0 r. \tag{8}$$

Equation (8) reveals that if the feedforward is parameterized in terms of the reference r, then the error in (8) can be made invariant under the choice of r, given that $F(\theta_j)$ is selected as $F(\theta_j) = P_0^{-1}$.

In order to retain the analytic solution to (6), the filter $F(\theta_j)$ is typically chosen as a polynomial function that is linear in θ_j, e.g., an FIR filter, see [6], [7]. Hence, for this particular choice, the basis functions are referred to as polynomial basis functions.

Consequently, $F(\theta_j) = P_0^{-1}$, can only be achieved if P_0 is restricted to be a rational function with a unit numerator, i.e., no zeros. In case this condition is violated, the achievable performance and extrapolation properties of ILC are severely deteriorated. Since typical physical systems are modeled using rational models that contain both poles and zeros, a unit numerator imposes a significant restriction.

D. Paper Contribution: ILC With Rational Basis Functions for Enhancing Performance and Extrapolation Properties

In view of the limitations imposed by the polynomial basis functions in Section III-C, this brief aims to investigate more general parameterizations that enhance: 1) tracking performance and 2) extrapolation properties of the learned feedforward command signal. This brief contains the following contributions:

1) general rational basis functions are proposed;
2) the consequences of a more general rational parameterization on the resulting ILC optimization problem are investigated, revealing a significantly more complex optimization problem;
3) a solution strategy is proposed to deal with the more difficult optimization problem that is introduced by the general rational parameterization;
4) the results are experimentally validated on a benchmark motion system and compared with preexisting results.

Preliminary research related to 1) and 3) appeared in [24]. This brief extends these initial findings with more theory and explanations, and includes an experimental validation.

IV. New Framework for ILC With Rational Basis Functions

In this section, the main contribution of this brief is presented: the formulation, analysis, and synthesis of an optimal ILC with rational basis functions, i.e., aspects 1)–3) in Section III-D. As is argued in Section III-C, the motivation for using such basis functions stems from (8), which reveals that parameterizing the feedforward command signal in terms of the reference signal enables extrapolation of the learned feedforward command signal to other reference trajectories.

Definition 2 (Rational Basis for Optimal ILC): The rational basis functions are defined as

$$f_j = F(\theta_j) r \tag{9}$$

where $F \in \mathcal{F}$

$$\mathcal{F} = \left\{ B(\theta_j)^{-1} A(\theta_j) \,\middle|\, \theta_j \in \mathbb{R}^{m_a + m_b} \right\} \tag{10}$$

and

$$A(\theta_j) = \sum_{i=1}^{m_a} \xi_i^A \theta_j[i]$$
$$B(\theta_j) = I + \sum_{i=1}^{m_b} \xi_i^B \theta_j[i + m_a].$$

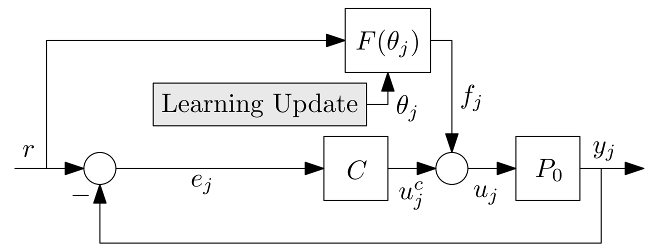

Fig. 2. Controller structure with rational basis.

Here, ξ_i^A and ξ_i^B are the convolution matrices corresponding to user-chosen polynomial transfer functions $\boldsymbol{\xi}_i^A(z)$ and $\boldsymbol{\xi}_i^B(z)$, respectively. The matrix $B(\theta_j)^{-1}$ is the convolution matrix corresponding to $\boldsymbol{B}(\theta_j, z)^{-1}$ with $\boldsymbol{B}(\theta_j, z) = 1 + \sum_{i=1}^{m_b} \boldsymbol{\xi}_i^B(z)\theta_j[i + m_a]$. The parameters $\theta_j = [\theta_j^A, \theta_j^B]^T$.

The ILC command f_j for ILC with rational basis functions in Definition 2 is implemented in the ILC setup of Fig. 1, see Fig. 2 for the resulting block diagram.

Remark 1: The underlying transfer function $\boldsymbol{F}(\theta_j, z)$ of $F(\theta_j)$ can be computed and analyzed in the frequency domain, e.g., by its frequency response function $\boldsymbol{F}(\theta_j, e^{i\omega})$, using: $\boldsymbol{A}(\theta_j, z) = \sum_{i=1}^{m_a} \boldsymbol{\xi}_i^A(z)\theta_j[i]$, $\boldsymbol{B}(\theta_j, z) = 1 + \sum_{i=1}^{m_b} \boldsymbol{\xi}_i^B(z)\theta_j[i + m_a]$, and $\boldsymbol{F}(\theta_j, z) = \boldsymbol{B}(\theta_j, z)^{-1}\boldsymbol{A}(\theta_j, z)$.

Remark 2: The classical ILC with polynomial basis functions approach, see Section III-C and [6], [7], is recovered by setting $m_b = 0$. Indeed, this leads to $B(\theta_j) = I$, and hence $F(\theta_j) = A(\theta_j)$ with A linear in θ_j.

A. Analysis of the Resulting Optimization Problem

Aspect 2) in Section III-D is elaborated on in this section. The difficulty associated with the rational basis function parameterization (10) involves the complexity of solving the corresponding optimization problem. In fact, the rational basis (10) in general prevents an analytic solution to (6). This is revealed by the following theorem.

Theorem 1: Let $W_f = W_{\Delta f} = 0$ and consider the parameterization (9) and (10). Then, $\mathcal{J}(f_{j+1})$, see (5), is nonlinear in θ_{j+1}.

Proof: Substitution of (9) and (10) into (4) yields

$$e_{j+1} = e_j + P_0 S_0 f_j - B(\theta_{j+1})^{-1} A(\theta_{j+1}) P_0 S_0 r.$$

Substituting e_{j+1} in (5)

$$\begin{aligned}\mathcal{J}(\theta_{j+1}) = {} & e_j^T W_e e_j + f_j^T S_0^T P_0^T W_e P_0 S_0 f_j + r^T S_0^T P_0^T B(\theta_{j+1})^{-T} \\ & \times A(\theta_{j+1}) W_e B(\theta_{j+1})^{-1} A(\theta_{j+1}) P_0 S_0 r \\ & + 2 e_j^T W_e P_0 S_0 f_j - 2 e_j^T W_e B(\theta_{j+1})^{-1} A(\theta_{j+1}) P_0 S_0 r \\ & - 2 f_j^T S_0^T P_0^T W_e B(\theta_{j+1})^{-1} A(\theta_{j+1}) P_0 S_0 r. \end{aligned} \tag{11}$$

■

Theorem 1 reveals that $B(\theta_{j+1})^{-1}$ in (11) leads to a performance criterion (5) that is nonlinear in the parameters θ_{j+1}. As a result, no analytic solution is available in general and the performance criterion is typically nonconvex in θ_{j+1}. In Section IV-B, an iterative solution is proposed to calculate θ_{j+1}, constituting contribution 3), see Section III-D.

B. Synthesis of Optimal ILC With Rational Basis Functions

In this section, an ILC algorithm is developed that enables optimal controller synthesis using rational basis functions. The main idea is to solve a sequence of least-squares problems and to consider the nonlinear terms as *a priori* unknown weighting functions. The basic concept is to recast (5) to

$$\mathcal{J}(\theta_{j+1}) = \|B(\theta_{j+1})^{-1}[B(\theta_{j+1})e_{j+1}]\|_{W_e} + \|B(\theta_{j+1})^{-1}[B(\theta_{j+1})f_{j+1}]\|_{W_f} + \|B(\theta_{j+1})^{-1}[B(\theta_{j+1})f_{j+1}] - f_j\|_{W_{\Delta f}}. \quad (12)$$

In (12), $\mathcal{J}(\theta_{j+1})$ is nonlinear in θ_{j+1} due to the term $B(\theta_{j+1})^{-1}$. However, $\mathcal{J}(\theta_{j+1})$ is linear in θ_{j+1} in the terms $B(\theta_{j+1})e_{j+1}$ and $B(\theta_{j+1})f_{j+1}$. In view of this distinction, an auxiliary index k is introduced, i.e., $\theta_{j+1}^{<k>}$ and $\theta_{j+1}^{<k-1>}$. These are substituted into (12), yielding

$$\mathcal{J}_k(\theta_{j+1}^{<k>}) = \left\| B\left(\theta_{j+1}^{<k-1>}\right)^{-1}\left[B\left(\theta_{j+1}^{<k>}\right)e_{j+1}^{<k>}\right]\right\|_{W_e} + \left\| B\left(\theta_{j+1}^{<k-1>}\right)^{-1}\left[B\left(\theta_{j+1}^{<k>}\right)f_{j+1}^{<k>}\right]\right\|_{W_f} + \left\| B\left(\theta_{j+1}^{<k-1>}\right)^{-1}\left[B\left(\theta_{j+1}^{<k>}\right)f_{j+1}^{<k>}\right] - f_j\right\|_{W_{\Delta f}} \quad (13)$$

where

$$e_{j+1}^{<k>} = e_j + P_0S_0f_j - B\left(\theta_{j+1}^{<k>}\right)^{-1}A\left(\theta_{j+1}^{<k>}\right)P_0S_0r.$$

Notice that (12) is recovered by setting $\theta_{j+1} = \theta_{j+1}^{<k>} = \theta_{j+1}^{<k-1>}$. In addition, notice that if $\theta_{j+1}^{<k-1>}$ is known, then $\mathcal{J}_k(\theta_{j+1}^{<k>})$ is a quadratic function of $\theta_{j+1}^{<k>}$. Consequently, $\theta_{j+1}^{<k>}$ can be calculated analytically. The basic principle is to fix the nonlinear $B\left(\theta_{j+1}^{<k-1>}\right)^{-1}$ at iteration k and interpret it as an iteratively adjusted weighting function. By iterating over $\theta_{j+1}^{<k>}$, it is aimed that the *a priori* unknown weighting by $B\left(\theta_{j+1}^{<k-1>}\right)^{-1}$ is effectively compensated after convergence of the iterative procedure. Clearly, this necessitates a solution to (13) for $\theta_{j+1}^{<k>}$, given $\theta_{j+1}^{<k-1>}$. The following theorem provides the analytic solution to $\theta_{j+1}^{<k>}$ that minimizes $\mathcal{J}_k(\theta_{j+1}^{<k>})$.

Theorem 2: Given $\theta_{j+1}^{<k-1>}$, f_j and e_j. Then, $\mathcal{J}_k(\theta_{j+1}^{<k>})$, see (13), is minimized by

$$\theta_{j+1}^{<k>} = L^{<k>}e_j + Q^{<k>}f_j \quad (14)$$

with

$$L^{<k>} = \left[\Psi_1^T W_e \Psi_1 + \Psi_2^T(W_f + W_{\Delta f})\Psi_2\right]^{-1}\Psi_1^T W_e B\left(\theta_{j+1}^{<k-1>}\right)^{-1}$$

$$Q^{<k>} = \left[\Psi_1^T W_e \Psi_1 + \Psi_2^T(W_f + W_{\Delta f})\Psi_2\right]^{-1} \times\left(\Psi_2^T W_{\Delta f} + \Psi_1^T W_e B\left(\theta_{j+1}^{<k-1>}\right)^{-1}PS\right)$$

where

$$\Psi_1 = B\left(\theta_{j+1}^{<k-1>}\right)^{-1}\left[\Psi_{PSr}^A, -\Psi_{e_j}^B - \Psi_{PSf_j}^B\right]$$

$$\Psi_2 = B\left(\theta_{j+1}^{<k-1>}\right)^{-1}\left[\Psi_r^A, 0\right].$$

Proof: Note that (13) is quadratic in $\theta_{j+1}^{<k>}$. A necessary condition for optimality is $\partial\mathcal{J}_k/\partial\theta_{j+1}^{<k>} = 0$. Solving this linear equation for $\theta_{j+1}^{<k>}$ yields the parameter update in (14). ■

Note that Theorem 2 in itself does not lead to the optimal solution of (6) in general, since $B\left(\theta_{j+1}^{<k-1>}\right)^{-1}$ is unknown. The proposed solution is to iteratively solve for $\theta_{j+1}^{<k>}$ in (13), given an $\theta_{j+1}^{<k=0>}$, for increasing k. In this approach, $B\left(\theta_{j+1}^{<k-1>}\right)^{-1}$ can be interpreted as an *a priori* unknown weighting in the cost function. This weighting is compensated for during each iteration over k by updating $L^{<k>}$ and $Q^{<k>}$. These steps are formulated in the following parameter update algorithm that addresses aspect 3), see Section III-D.

Algorithm 1: Given f_j and e_j, set $k = 0$ and initialize $\theta_{j+1}^{<k-1>} = \theta_j$. Then, perform the following sequence of steps.

1) Determine $L^{<k>}$, $Q^{<k>}$.
2) Determine $\theta_{j+1}^{<k+1>} = Q^{<k>}f_j + L^{<k>}e_j$.
3) Set $k \to k + 1$ and go back to (1) until an appropriate convergence condition is met: $\theta_{j+1}^{<k\to\infty>} = \theta_{j+1}$.

Theorem 2 and Algorithm 1 provide a new solution and algorithm to minimize $\mathcal{J}(\theta_{j+1})$ in (12), constituting the main result of this brief. The key novelty of these results lies in their use in optimal ILC algorithms. Indeed, related algorithms, see [16]–[18] are successfully used in system identification. Despite the fact that the objective function is nonconvex in general, practical use has revealed good convergence properties and in fact global convergence has been established for specific cases, see [25], [26].

Remark 3: In case Theorem 2 involves a nonminimum phase $\boldsymbol{B}(\theta_{j+1}^{<k-1>}, z)$ and hence unstable $\boldsymbol{B}(\theta_{j+1}^{<k-1>}, z)^{-1}$, the filtering operations cannot be performed in the usual manner, since time domain computation leads to unbounded results. Several approaches to calculate the filtered signals can be pursued, including: 1) approximations, see [15], [27], [28] and 2) exact methods, for instance, the stable inversion approach in [14]. In the latter, the filter is seen as a noncausal operator instead of an unstable one, see also [29, Sec. 1.5].

V. Connections to Preexisting Approaches

As mentioned throughout the preceding sections, specific basis choices can be recovered as special cases, including the results in [6], [7], and [13]. In this section, these specific parameterizations are compared with respect to the unified framework and solution as proposed in this brief.

A. FIR Structure

In [6], a polynomial basis $F(\theta_j)$ is used that is linear in the parameters θ_j. Given its close resemblance to the well-known FIR basis, the parameterization is referred to as FIR parameterization. This FIR parameterization connects to the rational basis functions in Definition 2 by setting $m_b = 0$ and $m_a = m$. As a result, $B(\theta_j) = I$. Furthermore, in [6], $W_e = I$ and $W_f = W_{\Delta f} = 0$. Finally, let $\boldsymbol{\xi}(z) = 1 - z^{-1}$, then the basis functions used are $\xi_1^A = \xi, \xi_2^A = \xi^2, \ldots, \xi_m^A = \xi^m$.

B. Extended FIR

In [13], a more general polynomial basis is presented that extends the FIR parameterization in [6]. By exploiting the commutativity property for SISO LTI systems, the framework in [13] can be recast in the form of Fig. 2 by selecting: $W_e = I, W_f = W_{\Delta f} = 0$; let $\boldsymbol{\xi}(z) = 1 - z^{-1}$, then the basis functions used are: $\xi_1^A = \xi$, $\xi_2^A = \xi^2, \ldots, \xi_{m_a}^A = \xi^{m_a}$, $\xi_1^B = \xi, \xi_2^B = \xi^2, \ldots, \xi_{m_b}^B = \xi^{m_b}$, and $B(\theta_{j+1})^{-1} := I$. This approach coincides with Algorithm 1, but without step 3. Hence, there is no compensation of the *a priori* unknown weighting function. As a result, an *a priori* unknown weighting $B(\theta_{j+1})$ is introduced in the performance criterion $\mathcal{J}_{\text{ERR}}(\theta_{j+1}) = \|[B(\theta_{j+1})e_{j+1}]\|_{W_e}$.

VI. Experimental Results

In this section, the proposed algorithm is experimentally demonstrated, revealing the increased performance and extrapolation properties in comparison with preexisting results. In particular, the following approaches are compared:

1) the proposed norm-optimal ILC with rational basis functions;

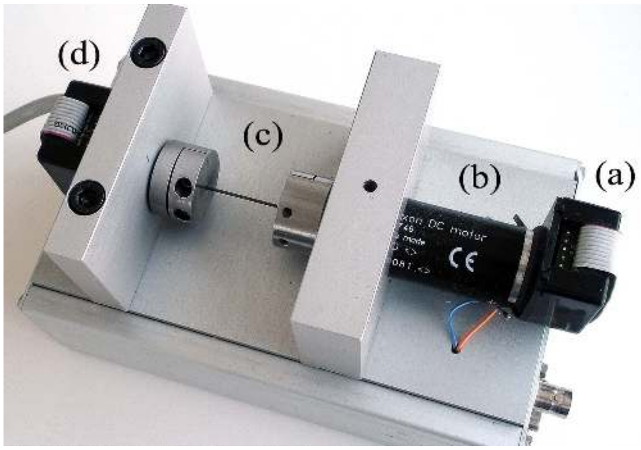

Fig. 3. Two-mass spring motion system. (a) Optical encoder. (b) Motor. (c) Mass-spring-mass. (d) Optical encoder.

2) the preexisting norm-optimal ILC with polynomial basis functions (FIR), see [6];
3) standard norm-optimal ILC.

A. Experimental Setup

The experimental two-mass spring motion system is shown in Fig. 3. The system consists of a current-controlled dc-motor [Fig. 3(b)] driving a mass (inertia) m_1 that is connected to mass (inertia) m_2 (Fig. 3) via a flexible shaft [Fig. 3(c)]. The positions of m_1 and m_2 are measured by optical encoders [Fig. 3(a) and (d)]. This system is well-suited to use as a benchmark system in order to examine prototype control algorithms [30]. In the results presented in this brief, only the position measurement of m_1 is used for control, and the position measurement of m_2 is ignored.

B. ILC Design

Step 1 (System Identification): Open-loop system identification is performed to identify the system P_0. The system is excited with random phased multisines at a sampling frequency f_s of 1 kHz. Both a parametric ($\boldsymbol{P}(z)$) and a nonparametric model (P_{frf}) is estimated using the measurement data. The corresponding Bode diagrams are shown in Fig. 4, accompanied with the 3σ (99.7%) confidence interval of the measured frequency response function P_{frf}.

Step 2 (Basis Functions and Weighting Matrices): The selection of basis functions ξ_i^A, ξ_I^B determines the model set $F(\theta)$, see Definition 2. Ideally, the structure of $A(\theta_j)$ and $B(\theta_j)$ should include the structure of the true inverse-system $P_0^{-1} = B_0(\theta)^{-1}A_0(\theta)$. Here, the structure of $A(\theta_j)$ and $B(\theta_j)$ is selected as follows: let $\boldsymbol{\xi}(z) = 1/T_s(1 - z^{-1})$ be a differentiator, then $\xi_1^A = \xi, \xi_2^A = \xi^2, \ldots, \xi_{m_a}^A = \xi^{m_a}$ and $\xi_1^B = \xi, \xi_2^B = \xi^2, \ldots, \xi_{m_b}^B = \xi^{m_b}$. This structure ensures the dc-gain $\boldsymbol{A}(\theta_j, z = 1) = 0$, and $\boldsymbol{B}(\theta_j, z = 1) = 1$, the latter guarantees that the rational structure F is well defined for all θ. This basis is expected to work well for all systems with infinite gain at zero frequency. If desired, it can easily be changed for systems with a finite dc-gain by adding a parameter such that $A(\theta_j, z = 1) = \theta_j$ [1]. In the FIR case, this basis selection allows θ_j^A to be directly interpreted as the feedforward parameters compensating for effects related to velocity, acceleration, jerk, and snap, see [12]. The scaling in $\boldsymbol{\xi}(z)$ with the sampling time T_s is to improve the numerical conditioning, and is related to the δ-operator approach in [31, Sec. 12.9].

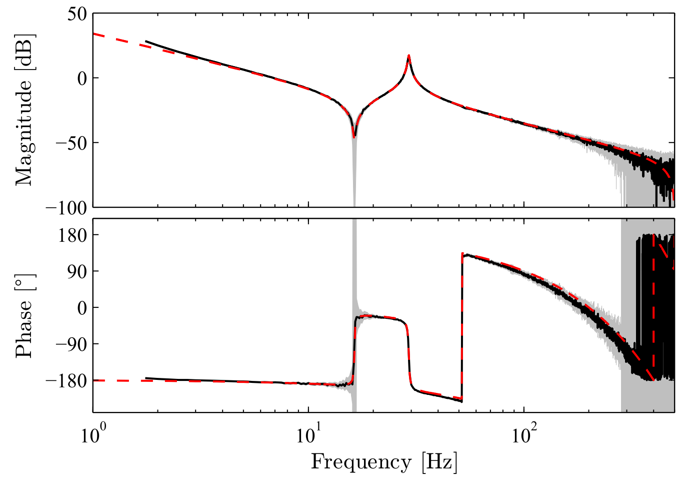

Fig. 4. Frequency response measurement P_{frf} (solid black line), 3σ confidence interval of P_{frf} (shaded gray area), and model $\boldsymbol{P}(e^{i\omega})$ (dashed red line).

Clearly, the optimal solution for the feedforward filter is given by $F(\theta) = P_0^{-1}$, since this choice leads to minimal $J(\theta)$. Using the identified model as a guideline, this choice corresponds to $m_a = 6$ and $m_b = 3$. However, to enable a fair comparison, also for the proposed rational basis, a restricted complexity parameterization is pursued such that under-modeling is present as follows.

1) Proposed $m_a = 4, m_b = 2$.
2) FIR $m_a = 4, m_b = 0$.

The rational filter is an extension of the FIR filter, where two zeros are added by setting $m_b = 2$. The selection of the number of parameters is similar to the problem of model order selection in system identification, see [32]. On the one hand, additional parameters increase the size of the model-set and therefore reduce bias, on the other hand, the variance on the parameters typically increases. Both aspects are a source of error between $F(\theta)$ and P_0^{-1}, and as such, manipulating this tradeoff is part of the controller design. Notice that standard norm-optimal ILC can be viewed as using an FIR structure with $m_a = n$ and $m_b = 0$ parameters in the ILC with basis functions framework. Hence, the basis functions affect the performance only and do not affect the convergence speed.

The weighting matrices in Definition 1 specify the performance and robustness objectives. In the results for both the FIR and rational structure presented in this brief, $W_e = I \cdot 10^3$, $W_f = W_{\Delta f} = 0$. This leads to an inverse model ILC, where the inverses in Theorem 2 are well-defined for the particular basis functions. For standard norm-optimal ILC, a small weighting on the learning speed is introduced: $W_{\Delta f} = I \cdot 0.05$.

C. Preliminary Simulation With a Fixed Reference

The proposed rational structure is compared with the FIR structure. The reference used is r_1, shown in Fig. 7. The corresponding ILC algorithms are invoked. To interpret the converged feedforward, the parameterized feedforward filters are visualized using a Bode diagram. The results are shown in Fig. 5, where $\boldsymbol{F}(\theta_\infty, e^{i\omega})^{-1}$ for the FIR and rational structures, and $\boldsymbol{P_0}(e^{i\omega})$ are compared. Fig. 5 shows that for frequencies up to 5 Hz the dynamics of P_0 are captured well by both approaches. The antiresonance (i.e., complex conjugates zeros) around 16.3 Hz is only captured by the proposed approach. The FIR structure does not have poles, hence its inverse does not have zeros and cannot accurately represent the antiresonance. Summarizing, from visual inspection it is concluded

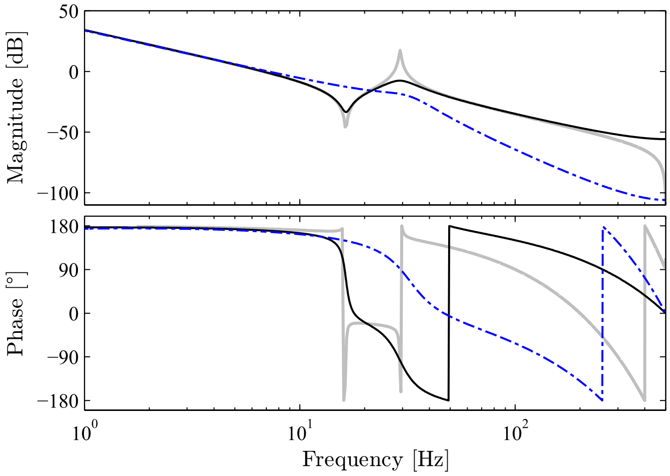

Fig. 5. Simulation results: $\boldsymbol{P_0}(e^{i\omega})$ and $\boldsymbol{F}(\theta_\infty, e^{i\omega})^{-1}$, $\boldsymbol{P_0}(e^{i\omega})$ (shaded gray area), proposed (solid black line), and FIR (dashed-dotted blue line).

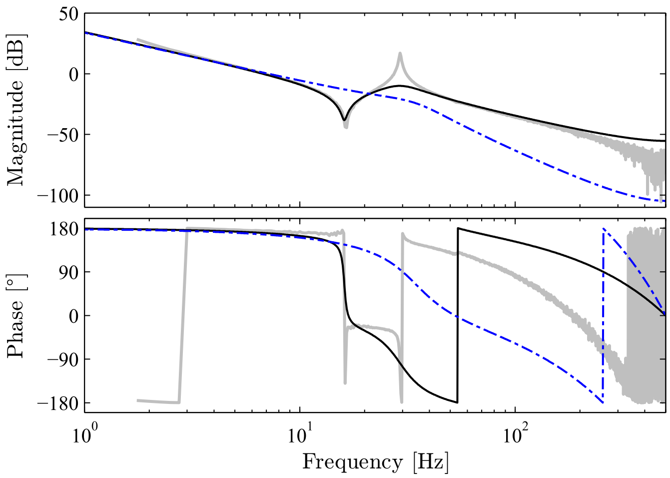

Fig. 6. Experimental results: P_{frf} and $\boldsymbol{F}(\theta_{14}, e^{i\omega})^{-1}$, P_{frf} (solid gray line), proposed (solid black line), and FIR (dashed-dotted blue line).

that for the proposed approach in this brief, $\boldsymbol{F}(\theta_\infty, e^{i\omega})^{-1}$ has the closest resemblance with the system $\boldsymbol{P_0}(e^{i\omega})$. It is therefore expected that the proposed approach has the best extrapolation capabilities if r changes, and this will be validated in Section VI-D in an experimental test case on the benchmark motion setup in Fig. 3.

D. Experimental Results

In this section, an experimental case study is presented where the extrapolation capabilities of the proposed rational and FIR feedforward parameterizations are verified and compared with standard norm-optimal ILC. The model, basis functions, and weighting matrices obtained in the previous section are used in the experiments. First, a preliminary experiment with reference r_1 establishes correspondence with the simulations followed by a case study on the extrapolation capabilities of the different approaches.

1) Experiment With a Fixed Reference: Fifteen trials are performed. The Bode diagram of the converged parameterized feedforward filters is shown in Fig. 6. Here, $\boldsymbol{F}(\theta_{14}, e^{i\omega})^{-1}$ for the FIR and rational structures and P_{frf} are compared. A visual comparison with the simulation, see Fig. 5, reveals similar results, except for the proposed approach, that shows a slightly better correspondence with P_{frf}, in particular at the antiresonance at 16.3 Hz.

2) Extrapolation of References: Three fourth order polynomial references are defined, see Fig. 7, where:

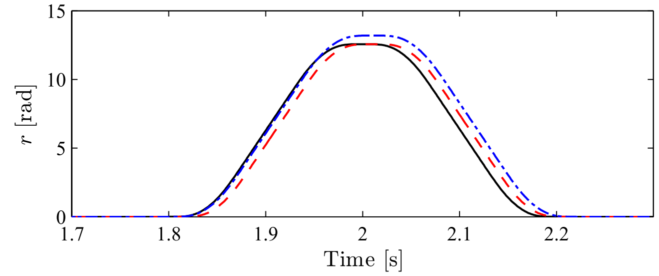

Fig. 7. Different references: r_1 (solid black line), r_2 (dashed red line), and r_3 (dashed-dotted blue line).

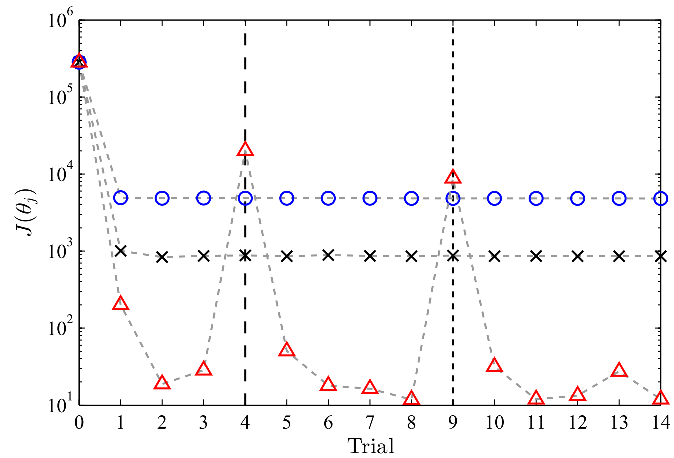

Fig. 8. Cost function values $J(\theta_j)$ where the proposed approach ($\times$) and the preexisting FIR approach ($\bigcirc$) are insensitive to the reference changes $r_1 \to r_2$ (black dashed line) at $j = 4$ and from $r_2 \to r_3$ at $j = 9$ (black dotted line), in contrast to standard norm-optimal ($\triangle$).

1) r_2 is equal to r_1 with 0.01 s delay;
2) r_3 has 5% extra distance with identical maximal velocity as r_1 and r_2.

In total, 15 trials are performed, where the reference is changed from r_1 to r_2 to r_3 at trials 4 and 9, respectively. The feedforward signal, or parameter vector θ_j, is not reinitialized when changing the reference in order to demonstrate the extrapolation capabilities.

The results are shown in Figs. 8 and 9. The cost function values, see Fig. 8, show that all approaches improve performance compared with feedback only ($f_0 = 0$). At $j = 4$ and $j = 9$, the reference is changed without reinitializing the feedforward signals. The key result is that both the FIR and proposed parameterization are insensitive to the change in the task, in contrast to standard norm-optimal ILC, where the reference changes results in a large increase in cost function value. The results in Fig. 8 also confirm that the proposed rational feedforward parameterization leads to improved tracking performance in comparison with the FIR structure. In addition, identical extrapolation capabilities are achieved.

The time domain tracking errors for trials 8 and 9 are shown in Fig. 9. These correspond with the time domain signals prior to the reference change $r_2 \to r_3$ (Fig. 9, left column) and after the change (Fig. 9, right column), respectively. These results confirm earlier conclusions, since: 1) the sensitivity of the tracking error when using standard norm-optimal ILC with respect to a small, i.e., 5% reference change is demonstrated; 2) the tracking performance of the FIR and proposed parameterizations are insensitive to the change in

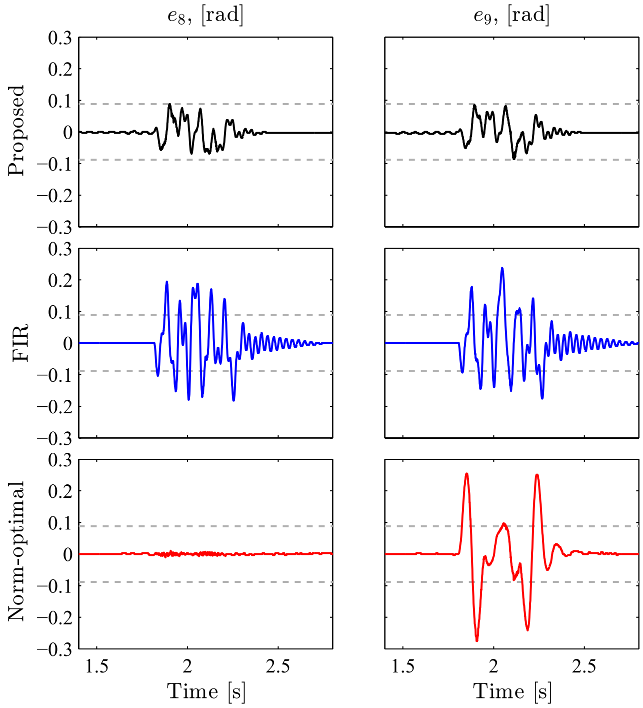

Fig. 9. Comparison of time domain tracking errors: prior to reference change (left column, $j = 8$), and after the reference change (right column, $j = 9$). Proposed approach (top row, black line), FIR (middle row, blue line), and norm-optimal (bottom row, red line).

reference; and 3) the proposed approach outperforms the approach with the FIR structure.

VII. Conclusion

In this brief, a novel framework for ILC with rational basis functions is presented. Herein, basis functions are adopted to enhance the extrapolation properties of learned command signals to other tasks. Indeed, in preexisting approaches, polynomial basis functions are employed that are only optimal for systems that have a unit numerator, i.e., no zeros.

The difficulty associated with rational basis functions lies in the synthesis of optimal ILC, since the analytic solution to optimal ILC algorithms is lost. In this brief, an iterative algorithm is proposed that effectively solves the optimization problem. It has close connections to well-known and powerful iterative solution methods in system identification.

The advantages of using rational basis functions in ILC are confirmed in a relevant experimental study: 1) improved performance with respect to preexisting methods that address extrapolation in ILC is demonstrated and 2) the performance is insensitive for the presented changes in the reference. The convergence aspects of the proposed method are experienced to be good, as is also confirmed in studies on related algorithms in system identification [18].

Ongoing research is toward: extending the approach to multiple-input multiple-output systems, investigating other solutions to the optimization problem, and addressing nonmeasurable performance variables, see [3], [33].

Acknowledgment

The authors would like to thank the late Prof. O. Bosgra, and Prof. M. Steinbuch, Eindhoven University of Technology for their fruitful discussions, guidance, and the contributions. The authors would also like to thank S. Koekebakker, Océ Technologies.

References

[1] K. Barton, D. Hoelzle, A. Alleyne, and A. Johnson, "Cross-coupled iterative learning control of systems with dissimilar dynamics: Design and implementation for manufacturing applications," *Int. J. Control*, vol. 84, no. 7, pp. 1223–1233, 2011.

[2] D. Hoelzle, A. Alleyne, and A. Johnson, "Basis task approach to iterative learning control with applications to micro-robotic deposition," *IEEE Trans. Control Syst. Technol.*, vol. 19, no. 5, pp. 1138–1148, Aug. 2011.

[3] J. Wallén, M. Norrlöf, and S. Gunnarsson, "A framework for analysis of observer-based ILC," *Asian J. Control*, vol. 13, no. 1, pp. 3–14, 2011.

[4] J. Bolder, B. Lemmen, S. Koekebakker, T. Oomen, O. Bosgra, and M. Steinbuch, "Iterative learning control with basis functions for media positioning in scanning inkjet printers," in *Proc. IEEE ISIC*, Oct. 2012, pp. 1255–1260.

[5] S. Mishra, J. Coaplen, and M. Tomizuka, "Precision positioning of wafer scanners segmented iterative learning control for nonrepetitive disturbances," *IEEE Control Syst.*, vol. 27, no. 4, pp. 20–25, Aug. 2007.

[6] S. van der Meulen, R. Tousain, and O. Bosgra, "Fixed structure feedforward controller design exploiting iterative trials: Application to a wafer stage and a desktop printer," *J. Dyn. Sys. Meas. Control*, vol. 130, no. 5, pp. 1–16, 2008.

[7] M. Heertjes, D. Hennekens, and M. Steinbuch, "MIMO feed-forward design in wafer scanners using a gradient approximation-based algorithm," *Control Eng. Pract.*, vol. 18, no. 5, pp. 495–506, 2010.

[8] S. Gunnarsson and M. Norrlöf, "On the design of ILC algorithms using optimization," *Automatica*, vol. 37, no. 12, pp. 2011–2016, Dec. 2001.

[9] J. van de Wijdeven and O. Bosgra, "Using basis functions in iterative learning control: Analysis and design theory," *Int. J. Control*, vol. 83, no. 4, pp. 661–675, 2010.

[10] M. Phan and J. Frueh, "Learning control for trajectory tracking using basis functions," in *Proc. IEEE 35th Decision Control Conf.*, vol. 3. Kobe, Japan, Dec. 1996, pp. 2490–2492.

[11] S. Oh, M. Phan, and R. Longman, "Use of decoupling basis functions in learning control for local learning and improved transients," *Adv. Astron. Sci.*, vol. 95, pp. 651–670, Nov. 1997.

[12] P. Lambrechts, M. Boerlage, and M. Steinbuch, "Trajectory planning and feedforward design for electromechanical motion systems," *Control Eng. Pract.*, vol. 13, no. 2, pp. 145–157, 2005.

[13] F. Boeren, D. Bruijnen, N. van Dijk, and T. Oomen. Joint input shaping and feedforward for point-to-point motion: Automated tuning for an industrial nanopositioning system. *Mechatronics* [Online]. Available: http://dx.doi.org/10.1016/j.mechatronics.2014.03.005

[14] S. Devasia, "Should model-based inverse inputs be used as feedforward under plant unvertainty?" *IEEE Trans. Autom. Control*, vol. 47, no. 11, pp. 1865–1871, Dec. 2002.

[15] J. A. Butterworth, L. Y. Pao, and D. Y. Abramovitch, "Analysis and comparison of three discrete-time feedforward model-based control techniques," *Mechatronics*, vol. 22, no. 5, pp. 577–587, Aug. 2012.

[16] K. Steiglitz and L. McBride, "A technique for the identification of linear systems," *IEEE Trans. Autom. Control*, vol. 10, no. 4, pp. 461–464, Oct. 1965.

[17] C. Sanathanan and J. Koerner, "Transfer function synthesis as a ratio of two complex polynomials," *IEEE Trans. Autom. Control*, vol. 8, no. 1, pp. 56–58, Jan. 1963.

[18] C. Bohn and H. Unbehauen, "Minmax and least squares multivariable transfer function curve fitting: Error criteria, algorithms and comparisons," in *Proc. Amer. Control Conf.*, vol. 5. Philadelphia, PA, USA, Jun. 1998, pp. 3189–3193.

[19] N. Amann, D. H. Owens, and E. Rogers, "Iterative learning control using optimal feedback and feedforward actions," *Int. J. Control*, vol. 65, no. 2, pp. 277–293, Feb. 1996.

[20] J. Lee, K. Lee, and W. Kim, "Model-based iterative learning control with a quadratic criterion for time-varying linear systems," *Automatica*, vol. 36, no. 5, pp. 641–657, May 2000.

[21] S. Mishra, U. Topcu, and M. Tomizuka, "Optimization-based constrained iterative learning control," *IEEE Trans. Control Syst. Technol.*, vol. 19, no. 6, pp. 1613–1621, Nov. 2011.

[22] P. Janssens, G. Pipeleers, and J. Swevers, "A data-driven constrained norm-optimal iterative learning control framework for LTI systems," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 2, pp. 546–551, Mar. 2013.

[23] D. A. Bristow, M. Tharayil, and A. G. Alleyne, "A survey of iterative learning control: A learning-based method for high-performance tracking control," *IEEE Control Syst. Mag.*, vol. 26, no. 3, pp. 96–114, Mar. 2006.

[24] J. Bolder, T. Oomen, and M. Steinbuch, "Exploiting rational basis functions in iterative learning control," in *Proc. IEEE 52nd Annu. CDC*, Florence, Italy, Dec. 2013, pp. 7321–7326.

[25] P. Stoica and T. Söderström, "The Steiglitz-McBride identification algorithm revisited–convergence analysis and accuracy aspects," *IEEE Trans. Autom. Control*, vol. 26, no. 3, pp. 712–717, Jun. 1981.

[26] P. Regalia, M. Mboup, and M. Ashari, "Existence of stationary points for pseudo-linear regression identification algorithms," *IEEE Trans. Autom. Control*, vol. 44, no. 5, pp. 994–998, May 1999.

[27] M. Tomizuka, "Zero phase error tracking algorithm for digital control," *J. Dyn. Syst., Meas., Control*, vol. 109, no. 1, pp. 65–68, Mar. 1987.

[28] E. Gross, M. Tomizuka, and W. Messner, "Cancellation of discrete time unstable zeros by feedforward control," *J. Dyn. Syst., Meas., Control, Trans. ASME*, vol. 116, no. 1, pp. 33–38, Mar. 1994.

[29] G. Vinnicombe, *Uncertainty and Feedback: $\mathcal{H}_\infty$ Loop-Shaping and the υ-Gap Metric*. London, U.K.: Imperial College Press, 2001.

[30] B. Wie and D. Bernstein, "Benchmark problems for robust control design," *J. Guid. Control Dyn.*, vol. 15, no. 5, pp. 1057–1059, Oct. 1992.

[31] G. Goodwin, S. Graebe, and M. Salgado, *Control System Design*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2000.

[32] T. Chen, H. Ohlsson, and L. Ljung, "On the estimation of transfer functions, regularizations and gaussian processes—Revisited," *Automatica*, vol. 48, no. 8, pp. 1525–1535, Aug. 2012.

[33] T. Oomen, J. van de Wijdeven, and O. Bosgra, "System identification and low-order optimal control of intersample behavior in ILC," *IEEE Trans. Autom. Control*, vol. 56, no. 11, pp. 2734–2739, Nov. 2011.