System Identification and Low-Order Optimal Control of Intersample Behavior in ILC

Tom Oomen, Jeroen van de Wijdeven, and Okko H. Bosgra

***Abstract*—Although iterative learning control (ILC) algorithms enable performance improvement for batch repetitive systems using limited system knowledge, at least an approximate model is essential. The aim of the present technical note is to develop an ILC framework for sampled-data systems, i.e., by incorporating the intersample response. Hereto, a novel parametric system identification procedure and a low-order optimal ILC controller synthesis procedure are presented that both incorporate the intersample behavior in a multirate framework. The results include i) improved computational properties compared to prior optimization-based ILC algorithms, and ii) improved performance of sampled-data systems compared to common discrete time ILC. These results are confirmed in a simulation example.**

***Index Terms*—Iterative learning control (ILC).**

I. Introduction

Iterative learning control (ILC) enables the performance improvement of batch repetitive systems based on measured data, requiring only limited model quality. ILC algorithms are generally implemented in a digital computer environment, thereby aiming at high performance at the sampling instants.

Pursuing only high at-sample performance goes at the expense of the intersample behavior in many relevant control situations, see [1]. Indeed, the intersample behavior is ignored in common ILC algorithms and in [2], [3] it is shown that several ILC algorithms actually deteriorate the performance of a sampled-data system instead of improving it. In [3], a multirate approach is presented that systematically deals with the intersample behavior in an optimal ILC framework and time-varying aspects introduced by the multirate problem setting are appropriately addressed.

Although the intersample behavior can be systematically dealt with in optimal ILC for sampled-data systems, existing solutions are limited to small-scale problems since the involved matrix dimensions and the associated computational complexity inflates rapidly for increasing problem sizes. Here, small-scale refers to i) small trial lengths, ii) small number of inputs and outputs [4], and iii) small amount of considered intersample data [3]. The inflating computational complexity originates from the use of convolution matrices, which essentially contain a non-parametric system model. The aim of this technical note is to develop more efficient ILC approaches for sampled-data systems. The key step in the developments is the use of parametric models. Consequently, the considered problem is decomposed into two subproblems: i) the construction of parametric models and ii) optimal ILC using parametric models. The key requirements for the approach are i) the intersample behavior should be addressed as motivated in [2], [3], ii) the algorithms should enable an efficient and numerically tractable implementation, and iii) the approach should be able to deal with the presence of a feedback controller.

Any ILC approach requires certain system knowledge reflected by a model, since large model errors can result in a divergent learning process as is evidenced by robustness analyses [5]. Although modeling for ILC is considered in [6], the approach is restricted to nonparametric models. In [7], [8], ILC for improved experiment design is considered, however, this does not deal with identification in view of ILC. Thus, although it has often been argued, e.g., [8], that the application should be taken into account when identifying models, the ILC application has remained largely unexplored, even in the simplified case where the intersample behavior is neglected.

When identifying models in view of ILC for sampled-data systems, the model should represent the intersample behavior. In general, the required models cannot be uniquely recovered from their discrete time model representation, as is shown in [9]. As a basic requirement, system identification should thus be able to deal with time-varying aspects that are generally introduced by incorporating the intersample behavior in a multirate or sampled-data setting. In the case of sampled-data systems, these aspects are commonly dealt with through lifting [10]. Key contributions include dealing with the causality constraint of the identified models, see [11], and dealing with non-uniformly sampled data, e.g., [12]. However, as shown in Section III, these causality problems and non-uniformly sampled data do not arise in the considered ILC problem.

In many optimization-based ILC algorithms, including [3], [4], [13], the use of convolution matrices leads to intractable computations for large-scale problems. Here, large-scale refers to the situation where either one of the variables regarding the trial length N^l, the number of inputs n_u and outputs n_y, or the amount of considered intersample data represented by F is large. Specifically, the typical matrix dimensions associated with the approaches in [3], [4], [13] are $(Fn_yN^l)\times(n_uN^l)$, leading to a typical efficiency of $4Fn_y(N^l)^3n_u^2+8n_u^3(N^l)^3$ flops. For a discrete time ILC application such as precision motion systems, see [1], 6 inputs, 6 outputs, sampling frequencies in the order of 10^4 Hz, and tasks in the order of several seconds already lead to intractable computations when using standard numerical algorithms. In contrast, it turns out that in the case that the underlying system is considered in its original dynamic form, a connection can be established between linear quadratic optimal control and ILC, leading to significantly more efficient algorithms. Results in this direction are presented in [14], [15], however, these approaches are restricted to a specific optimization criterion and cannot deal with the intersample behavior. The aim of the present technical note is to develop efficient ILC algorithms that deal with the intersample behavior.

At present, the system identification problem for sampled-data ILC and the optimal sampled-data ILC problem have not yet been appropriately dealt with. The key contribution of this technical note includes a low-order parametric approach to system identification for sampled-data ILC and an optimal ILC controller synthesis procedure. The proposed system identification procedure applies to closed-loop systems and delivers a model that encompasses the intersample behavior in a multirate setting. The resulting model can be used directly in the presented optimal ILC synthesis procedure. In contrast to pre-existing approaches based on convolution matrices, the involved computations involve a discrete time Riccati equation and the involved matrix dimen-

Manuscript received November 21, 2008; revised November 18, 2010; accepted June 02, 2011. Date of publication June 23, 2011; date of current version November 02, 2011. Recommended by Associate Editor T. Zhou.

T. Oomen is with the Department of Mechanical Engineering, Control Systems Technology Group, Eindhoven University of Technology, Eindhoven 5600 MB, The Netherlands (e-mail: t.a.e.oomen@tue.nl).

J. van de Wijdeven was with the Department of Mechanical Engineering, Control Systems Technology Group, Eindhoven University of Technology, Eindhoven 5600 MB, The Netherlands. He is now with TMC, Eindhoven 5600 AS, The Netherlands (e-mail: j.j.m.v.d.wijdeven@gmail.com).

O. H. Bosgra is with the Department of Mechanical Engineering, Control Systems Technology Group, Eindhoven University of Technology, Eindhoven 5600 MB, The Netherlands and is also with the Delft University of Technology, Delft 2628 CN, The Netherlands (e-mail: o.h.bosgra@tudelft.nl).

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

Digital Object Identifier 10.1109/TAC.2011.2160596

sions are bounded by the order of the underlying model, which is typically orders of magnitude smaller than the trial length and independent of the incorporation of intersample behavior and number of inputs and outputs. The low-order approach thus enables an efficient and numerically tractable implementation for large-scale problems. In addition, the use of low-order parametric models enables a reduction of non-systematic errors during system identification, e.g., in terms of variance. In the next section, the considered problem is defined in detail.

The technical note is organized as follows. In the next section, the problem formulation is defined precisely. Then, in Section III, the system identification for multirate ILC problem is addressed. A low-order solution to the optimal multirate ILC problem is presented in Section IV. The discrete time case is recovered for $F = 1$, where F is defined in the next section. Then, in Section V, an example illustrates the obtained results. Finally, conclusions are provided in Section VI.

Notation

$t \in \mathbb{Z}$ and $t_c \in \mathbb{R}$ denote discrete time and continuous time, respectively. The sampling frequency of a discrete time signal is indicated by a superscript, i.e., the signals x^h and x^l operate at the high sampling frequency f^h and low sampling frequency f^l, respectively. Sampling is assumed to be non-pathological. The forward shift operator is denoted by q. The delay operator D_τ is defined by $(D_\tau f)(t) = f(t-\tau)$, where $\tau \in t$ and G is time invariant if $D_\tau G = G D_\tau$, $\forall \tau \in t$. The truncation operator is defined by $(P_\tau f)(t) = f(t)$ for $t \leq \tau$ and $(P_\tau f)(t) = 0$ for $t > \tau$. The discrete Fourier transform of a signal x is given by $X_N(\omega) = (1/\sqrt{N}) \sum_{t=1}^{N} x(t) e^{j\omega t}$. The notation $\|x\|_W$ denotes $x^T W x$, whereas $\|x\|_p$ denotes the p-norm for $p \in \mathbb{N}$. The Kronecker product is denoted by $\otimes$.

II. Problem Formulation

Consider the setup in Fig. 1, where P denotes the continuous time plant with n_y outputs and n_u inputs. Also, continuous time signals are indicated by solid lines, whereas discrete time signals at sampling frequency f^l and f^h are indicated by dashed and dotted lines, respectively. Here, $f^h = 1/T^h = F f^l$, $F \in \mathbb{N}$, $t_i \in \mathbb{Z}$, and T^h denotes sampling time. The plant input u and error signal e are given by

$$\begin{aligned} u(t_c) &= \mathcal{H}^h \mathcal{H}_u (w^l(t) + C^{d,l} \mathcal{S}_d \mathcal{S}^h e(t_c)),\ tT^l \leq t_c < (t+1)T^l \\ e(t_c) &= r(t_c) - y(t_c) \end{aligned} \tag{1}$$

where $C^{d,l}$ is a discrete time feedback controller operating at sampling frequency $f^l = 1/T^l$. To convert continuous time signals into discrete time signals and *vice versa* in (1), the ideal sampler and zero-order-hold are defined by

$$\begin{aligned} \mathcal{S}^h &: e(t_c) \mapsto e^h(t),\ e^h(t_i) = e(t_i T^h) \\ \mathcal{H}^h &: u^h(t) \mapsto u(t_c),\ u(t_i T^h + \tau) = u^h(t_i),\ \tau = [0, T^h) \end{aligned}$$

respectively. To reduce the sampling frequency of a discrete time signal, see Fig. 1, the downsampling operator $\mathcal{S}_d$ is defined by

$$\mathcal{S}_d : e^h(t) \mapsto e^l(t),\ e^l(t_i) = e^h(F t_i), \quad t_i \in \mathbb{Z} \tag{2}$$

In addition, the multirate zero-order-hold $\mathcal{H}_u$ is defined as [1]

$$\mathcal{H}_u = \mathcal{I}^F(q) \mathcal{S}_u \tag{3}$$

where the upsampler $\mathcal{S}_u$ and zero-order-hold interpolator $\mathcal{I}^F(q)$ are given by

$$\begin{aligned} \mathcal{S}_u &: u^l(t) \mapsto x^h(t),\ x^h(t_i) := \begin{cases} u^l(\frac{t_i}{F}) & \text{for } t_i \in t,\ \frac{t_i}{F} \in \mathbb{Z} \\ 0 & \text{for } t_i \in t,\ \frac{t_i}{F} \notin \mathbb{Z} \end{cases} \\ \mathcal{I}^F(q) &= \sum_{f=0}^{F-1} q^{-f} \end{aligned} \tag{4}$$

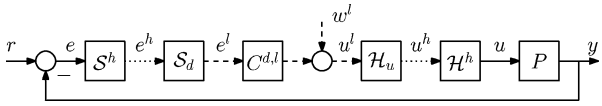

Fig. 1. Closed-loop sampled-data ILC setup.

respectively. Here, $\mathcal{S}_u$ increases the sampling frequency of a discrete time signal, whereas $\mathcal{I}^F(q)$ interpolates these values using a zero-order-hold interpolation scheme.

The motivation for considering two distinct sampling frequencies f^l and f^h stems from the constraint that the sampling frequency of the feedback controller is upper bounded, since the new control input has to be computed in real time. In contrast, in many control applications it is possible to measure and store the error signal at a higher sampling frequency f^h. As a result, the measured signals at the high sampling frequency can be processed in between consecutive trials by the ILC algorithm. Finally, it is remarked that the control signal w^l is assumed additive to the controller output, see (1). This assumption is nonrestrictive and stems from common ILC implementations [3]. Taking into account these aspects leads to the optimal multirate ILC problem.

Problem 1 (Optimal Multirate ILC): Given the criterion $\mathcal{J}_{\mathrm{MR}}(w^l, e^h)$, determine

$$w^l_{\mathrm{MR}\star} = \arg\min_{w^l} \mathcal{J}_{\mathrm{MR}}(w^l, e^h).$$

Here, $\mathcal{J}_{\mathrm{MR}}(w^l, e^h)$ is a norm-based criterion, e.g., $\mathcal{J}_{\mathrm{MR}}(w^l, e^h) = \left\| \begin{bmatrix} (w^l)^T & (e^h)^T \end{bmatrix}^T \right\|_2^2$. The multirate ILC problem approximates the sampled-data ILC problem that is introduced in [3]. In fact, convergence to the underlying sampled-data problem can be established under certain technical conditions. In contrast, common ILC approaches are implemented in discrete time and involve a criterion of the form $\mathcal{J}_{\mathrm{DT}}(w^l, e^l)$, e.g., $\mathcal{J}_{\mathrm{DT}}(w^l, e^l) = \left\| \begin{bmatrix} (w^l)^T & (e^l)^T \end{bmatrix}^T \right\|_2^2$. Thus, discrete time ILC approaches neglect the intersample behavior since e^l is considered instead of e^h.

The basic idea in ILC is to iterate over k, where at iteration k measurements $e^h_{\langle k \rangle}(t)$, $t \in [0, N^h - 1]$ and $w^l_{\langle k \rangle}(t)$, $t \in [0, N^l - 1]$ are given. Then, based on these measurements, $w^l_{\langle k+1 \rangle}(t)$ is determined such that $\mathcal{J}_{\mathrm{MR}\langle k+1 \rangle}$ is minimized. If convergent, $w^l_{\langle k \rangle} \to w^l_{\mathrm{MR}\star}$ for $k \to \infty$. ILC algorithms require a model that approximates the underlying system. Specifically, a model of the operator $J_{\mathrm{MR}} : w^l \mapsto -e^h$ is required for the optimal multirate ILC problem, leading to the following system identification problem.

Problem 2 (System Identification for Multirate ILC): Given the closed-loop setup of Fig. 1, identify a model $J_{\mathrm{MR}} : w^l \mapsto -e^h$ that is suitable for solving the optimal multirate ILC problem.

Standard optimal ILC algorithms, including [13], and system identification approaches, including [16], do not solve Problem 1 and Problem 2, since these approaches typically require time-invariance of the system. Indeed, the system J_{MR} is generally not time-invariant, even if both P and $C^{d,l}$ are time-invariant. Time-variance of J_{MR} for $F > 1$ can be understood if it is observed that e^h operates at a higher sampling frequency than w^l.

Although a solution to Problem 1 is proposed in [3], the involved matrix dimensions inflate for increasing n_y, n_u, f^l, F. This precludes the solution in [3] to large-scale ILC problems that are commonly encountered in applications. In addition, the system identification for multirate ILC problem has not been dealt with in the literature.

III. System Identification for Multirate ILC

Time-variance of multirate operators obstructs the use of standard system identification and optimization techniques to address Problem

1 and Problem 2. Throughout, the lifting operator $\mathcal{L}$, see [10], has a central role in the derivations and is defined by

$$\widetilde{e} = \mathcal{L}e^h, \text{ where } \widetilde{e}(t_i) = \begin{bmatrix} e^h(Ft_i) \\ e^h(Ft_i+1) \\ \vdots \\ e^h(Ft_i+F-1) \end{bmatrix}. \tag{5}$$

Here, e^h is an n_y dimensional signal with sampling frequency f^h, whereas $\widetilde{e}$ is an Fn_y dimensional signal with sampling frequency f^l. The following properties of the lifting operator are relevant for subsequent developments.

Lemma 3: The lifting operator (5)

a) is norm-preserving, since $\|\widetilde{e}\|_p = \|e^h\|_p$.
b) is not a causal operator for $F > 1$.

Proof: Property (a) follows along similar lines as in [10, Page 420]. To prove Property b, causality requires $P_\tau G P_\tau = P_\tau G, \forall \tau \in t$, which does not hold if $\tau = 0$ and $F > 1$. ∎

The following theorem is essential for subsequent developments. Here, $P^{d,h} = \mathcal{S}^h P \mathcal{H}^h$ and $P^{d,l} = \mathcal{S}_d P^{d,h} \mathcal{H}_u$ are LTI systems operating with sampling time T^h and T^l, respectively. In addition, $\widetilde{J_{\mathrm{MR}}} = \mathcal{L} J_{\mathrm{MR}}$.

Theorem 4: Let $(A_p^h, B_p^h, C_p^h, D_p^h)$ be a state-space realization of $P^{d,h}$, let $(A_c^l, B_c^l, C_c^l, D_c^l)$ be a state-space realization for $C^{d,l}$, and assume that $D_p^h D_c^l = 0$, $D_c^l D_p^h = 0$. Then, a state-space realization for $\widetilde{J_{\mathrm{MR}}} = \mathcal{L}P^{d,h}\mathcal{H}_u(I + C^{d,l}P^{d,l})^{-1}$ is given by $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$, where

$$\widetilde{A} = \begin{bmatrix} (A_p^h)^F - \mathfrak{B} D_c^l C_p^h & \mathfrak{B} C_c^l \\ -B_c^l C_p^h & A_c^l - B_c^l D_p^h C_c^l \end{bmatrix} \quad \widetilde{B} = \begin{bmatrix} \mathfrak{B} \\ -B_c^l D_p^h \end{bmatrix}$$

$$\widetilde{C} = \begin{bmatrix} C_p^h & D_p^h C_c^l \\ C_p^h A_p^h - C_p^h B_p^h D_c^l C_p^h & D_p^h C_c^l + C_p^h B_p^h C_c^l \\ \vdots & \vdots \\ C_p^h (A_p^h)^{F-1} - \mathfrak{D} D_c^l C_p^h & \mathfrak{D} C_c^l \end{bmatrix}$$

$$\widetilde{D} = \begin{bmatrix} D_p^h \\ D_p^h + C_p^h B_p^h \\ \vdots \\ \mathfrak{D} \end{bmatrix}$$

$$\mathfrak{B} = \sum_{f=0}^{F-1} (A_p^h)^f B_p^h, \quad \mathfrak{D} = D_p^h + \sum_{f=0}^{F-2} C_p^h (A_p^h)^f B_p^h.$$

Proof: Let x_p^h denote the state vector of $P^{d,h}$. Using (3) and successive substitution yields

$$x_p^h(Ft_i + q) = (A_p^h)^q x_p^h(Ft_i) + \mathfrak{B} u^h(Ft_i), q \in \mathbb{N}, 1 \le q \le F.$$

Subsequent application of (4) and (2) yields

$$x_p^l(t_i + 1) = (A_p^h)^F x_p^l(t_i) + \mathfrak{B}(u_c^l(t_i) + w^l(t_i)) \tag{6}$$

and corresponding output equation

$$y^l(t_i) = C_p^h x^l(t_i) + D_p^h(u_c^l(t_i) + w^l(t_i)). \tag{7}$$

Next, substitution of (6) and (7) into the controller equations $x_c^l(t_i + 1) = A_c^l x_c^l(t_i) - B_c^l y^l(t_i)$ and $u_c^l(t_i) = C_c^l x_c^l(t_i) - D_c^l y^l(t_i)$, respectively, yields the required matrices $\widetilde{A}$ and $\widetilde{B}$. Next, using (3) reveals that $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ indeed is a state-space realization for the lifted system. ∎

The assumption $D_p^h D_c^l = 0$, $D_c^l D_p^h = 0$ is a sufficient condition for well-posedness. A derivation of Theorem 4 without the assumption $D_p^h D_c^l = 0$, $D_c^l D_p^h = 0$ is straightforward and leads to more complicated state-space matrices. Theorem 4 leads to the following result that is required to enable application of system identification tools for finite dimensional LTI systems.

Theorem 5: Let $P^{d,h}$ be internally stabilized by $C^{d,l}$ with minimal state space realizations $(A_p^h, B_p^h, C_p^h, D_p^h)$ and $(A_c^l, B_c^l, C_c^l, D_c^l)$, respectively. Then,

a) $\widetilde{J_{\mathrm{MR}}}$ is LTI;
b) $\widetilde{J_{\mathrm{MR}}}$ has McMillan degree upper bounded by the sum of the McMillan degrees of P and $C^{d,l}$;
c) $\widetilde{J_{\mathrm{MR}}} \in \mathcal{RH}_\infty^{Fn_y \times n_u}$.

Proof: Observe that

$$D_\tau(\mathcal{L}P^{d,h}\mathcal{H}_u(I + C^{d,l}P^{d,l})^{-1}) = \mathcal{L}P^{d,h}\mathcal{H}_u(I + C^{d,l}P^{d,l})^{-1}D_\tau \ \forall \tau \in t$$

which proves a). To show b), note that the McMillan degree of a proper system equals the state dimension of a minimal state-space realization. Next, the state-space realization of the system in Theorem 4 has a McMillan degree which is equal to the sum of the McMillan degrees of $C^{d,l}$ and P, where the McMillan degree of P is invariant under (down)sampling, see [3]. Finally, inequality is obtained by possible nonminimality of $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$, which can be caused by pole/zero cancellations between $C^{d,l}$ and $P^{d,l}$. Part c) follows from the fact that $C^{d,l}$ is internally stabilizing, hence $\widetilde{A}$ is strictly Schur and as a result $\mathcal{L}P^{d,h}\mathcal{H}_u(I + C^{d,l}P^{d,l})^{-1} \in \mathcal{RH}_\infty^{Fn_y \times n_u}$. ∎

The results of Theorem 5 enable the solution of Problem 2. The following procedure is suggested.

Procedure 6: Consider the setup in Fig. 1. Then,

i) set $r = 0$, apply an excitation signal w^l and measure the output e^h;
ii) lift: $\widetilde{e} = \mathcal{L}(e^h)$;
iii) identify a model $\widehat{\widetilde{J_{\mathrm{MR}}}}$ with input w^l and output $-\widetilde{e}$.

The following remarks are relevant. Firstly, although $\widetilde{J_{\mathrm{MR}}}$ is a closed-loop system, the identification problem is posed in an open-loop setting and no problems are introduced by the closed-loop operation of the system. In fact, the proposed approach is closely related to the indirect approach to closed-loop system identification as is discussed in [16, Section 13.5]. Secondly, in view of Theorem 5 , the operator $\widetilde{J_{\mathrm{MR}}}$ is LTI, enabling the use of standard system identification techniques. In fact, in view of Lemma 3 a), the identification criterion in many cases is invariant under lifting, hence an identified model that is optimal in the lifted domain is generally also optimal in the physical time domain. Thirdly, note that Step 3 in Procedure 6 involves lifting, which is a non-causal operation. In contrast to, e.g., [11], where related but different approaches to the identification of multirate systems are presented, no additional constraints have to be imposed in the Procedure 6 to avoid non-causality of the underlying models in the physical time domain. This is due to the lack of an inverse lifting operator in the definition of $\widetilde{J_{\mathrm{MR}}}$. Finally, it is remarked that a model structure and order have to be selected in Procedure 6 Item iii). Such a model structure selection problem is inherent in any identification problem, see, e.g., [16, Chapter 16] for a discussion of this aspect. Any order selection procedure for open-loop systems can be used for the considered problem in this technical note. The order selection procedure generally leads to a McMillan degree that is bounded by the sum of the McMillan degrees of P and $C^{d,l}$.

Next, it is shown that the model $\widehat{\widetilde{J_{\mathrm{MR}}}}$ enables the design of low-order optimal ILC controllers, see Problem 1.

IV. Optimal Multirate ILC

In this section, a low-order state-space solution is presented to Problem 1. The following criterion is considered, which generalizes optimal ILC to incorporate intersample behavior.

Definition 7: The criterion for determining the optimal command input for trial $k+1$, i.e., $w^l_{\langle k+1\rangle}$, is defined as

$$\mathcal{J}_{\mathrm{MR}\,\langle k+1\rangle} = \frac{1}{2}\sum_{t=0}^{N^h-1}\left[\|e^h_{\langle k+1\rangle}(t)\|_{W_e}\right] + \frac{1}{2}\sum_{t=0}^{N^l-1}\left[\|w^l{}_{\langle k+1\rangle}(t)\|_{W_w} + \|w^l{}_{\langle k+1\rangle}(t) - w^l{}_{\langle k\rangle}(t)\|_{W_\Delta}\right]. \quad (8)$$

In (8), N^l and N^h denote the trial length at sampling frequency f^l and f^h, respectively. Hence, a finite time task is considered at each iteration. Additionally, W_e, W_w, and W_Δ are time-dependent weighting matrices, see [3] for guidelines on selecting these design parameters. The key issue in (8) is that the signals are defined in distinct function spaces to address the intersample behavior. The following lemma reveals that the lifting of signals can be exploited to recast the criterion (8) into a form that facilitates the computation of optimal ILC controllers.

Lemma 8: Consider the lifting operator (5) and criterion (8) in Definition 7. Then

$$\mathcal{J}_{\mathrm{MR}\,\langle k+1\rangle} = \frac{1}{2}\sum_{t=0}^{N^l-1}\left[\|\widetilde{e}_{\langle k+1\rangle}\|_{\widetilde{W_e}} + \|w^l{}_{\langle k+1\rangle}\|_{W_w} + \|w^l{}_{\langle k+1\rangle} - w^l{}_{\langle k\rangle}\|_{W_\Delta}\right] \quad (9)$$

where $\widetilde{W_e} := I_F \otimes W_e$.

Proof: Follows by applying (5) to (8). ∎

Lemma 8 enables the direct use of the identified LTI model in Section III to solve the multirate ILC problem, thereby explicitly addressing the intersample behavior. The solution to the multirate ILC problem via the lifting operator is the main result of this section and is given in the following theorem.

Theorem 9: Consider the criterion (9) and model $\widetilde{J}_{\mathrm{MR}} = (\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$. Then

$$w^l_{\langle k+1\rangle}(t) = C^{\mathrm{ILC}}\,[\,\widetilde{e}_{\langle k\rangle}(t)^T \quad w^l_{\langle k\rangle}(t)^T \quad g_{\langle k\rangle}(t+1)^T\,]^T \quad (10)$$

where C^{ILC} is given by the state-space realization

$$C^{\mathrm{ILC}} = \left[\begin{array}{c|ccc} \widetilde{A}-\widetilde{B}L(t) & \widetilde{B}L_e(t) & -\widetilde{B}L_w(t) & \widetilde{B}L_g(t) \\ \hline -L(t) & L_e(t) & I-L_w(t) & L_g(t) \end{array}\right]$$

$$L(t) = \Gamma_2^{-1}(\widetilde{B}^T P(t+1)A + \widetilde{D}W_e\widetilde{C}),$$
$$L_e(t) = \Gamma_2^{-1}, \widetilde{D}^T W_e,$$
$$L_w(t) = \Gamma_2^{-1}W_w, \qquad L_g(t) = \Gamma_2^{-1}\widetilde{B}^T$$
$$\Gamma_1 = W_w + W_\Delta + \widetilde{D}^T W_e \widetilde{D}, \quad \Gamma_2 = \Gamma_1 + \widetilde{B}^T P(t+1)\widetilde{B}$$

and P and $g_{\langle k\rangle}$ are given by the backward recursions

$$P(t) = H_{21} + H_{22}P(t+1)(I - H_{12}P(t+1))^{-1}H_{11} \quad (11)$$

$$g_{\langle k\rangle}(t) = (H_{22} + H_{22}P(t+1)(I - H_{12}P(t+1))^{-1}H_{12})g_{\langle k\rangle}(t+1) - (E_{21} + H_{22}P(t+1)(I - H_{12}P(t+1))^{-1}E_{11}\widetilde{e}_{\langle k\rangle}(t) - (E_{22} + H_{22}P(t+1)(I - H_{12}P(t+1))^{-1}E_{12}w^l_{\langle k\rangle}(t) \quad (12)$$

where $P(N^l) = 0$, $g_{\langle k\rangle}(N^l) = 0$, and

$$H_{11} = \widetilde{A} - \widetilde{B}\Gamma_1^{-1}\widetilde{D}^T\widetilde{W}_e\widetilde{C} \quad H_{12} = -\widetilde{B}\Gamma_1^{-1}\widetilde{B}^T$$
$$H_{21} = \widetilde{C}^T\widetilde{W}_e\widetilde{C} - \widetilde{C}^T\widetilde{W}_e\widetilde{D}\Gamma_1^{-1}\widetilde{D}^T\widetilde{W}_e\widetilde{C}$$
$$H_{22} = \widetilde{A}^T - \widetilde{C}^T\widetilde{W}_e\widetilde{D}\Gamma_1^{-1}\widetilde{B}^T$$
$$E_{11} = \widetilde{B}\Gamma_1^{-1}\widetilde{D}^T\widetilde{W}_e \qquad E_{12} = -\widetilde{B}\Gamma_1^{-1}W_w$$
$$E_{21} = -\widetilde{C}^T\widetilde{W}_e + \widetilde{C}^T W_e\widetilde{D}\Gamma_1^{-1}\widetilde{D}^T\widetilde{W}_e$$
$$E_{22} = -\widetilde{C}^T\widetilde{W}_e\widetilde{D}\Gamma_1^{-1}W_w$$

Proof: First, define

$$\Delta w^l_{\langle k\rangle} = w^l_{\langle k+1\rangle} - w^l_{\langle k\rangle} \quad (13)$$

and use $\widetilde{e}_{\langle k+1\rangle} = \widetilde{e}_{\langle k\rangle} - \widetilde{J_{\mathrm{MR}}}\Delta w^l_{\langle k\rangle}$ to rewrite (9) as

$$\mathcal{J}_{\mathrm{MR}\,\langle k+1\rangle} = \frac{1}{2}\sum_{t=0}^{N^l-1}\Big[\|\widetilde{C}\Delta x_k + \widetilde{D}\Delta w^l_k - \widetilde{e}_k\|_{\widetilde{W_e}} + \|\Delta w^l_k + w^l_k\|_{W_w} + \|\Delta w^l_k\|_{W_\Delta}\Big], \quad (14)$$

$$\text{subject to} \quad \Delta x_{\langle k+1\rangle} = \widetilde{A}\Delta x_{\langle k\rangle} + \widetilde{B}\Delta w^l_{\langle k\rangle}. \quad (15)$$

Consider the Hamiltonian function

$$\mathcal{H}_{\langle k\rangle} = \|\widetilde{C}\Delta x_k(t) + \widetilde{D}\Delta w^l_k(t) - \widetilde{e}_k(t)\|_{\widetilde{W_e}} + \|\Delta w^l_k(t) + w^l_k(t)\|_{W_w} + \|\Delta w^l_k(t)\|_{W_\Delta} + \lambda^T_{\langle k\rangle}(t+1)\left(\widetilde{A}\Delta x_{\langle k\rangle}(t) + \widetilde{B}\Delta w^l_{\langle k\rangle}(t)\right). \quad (16)$$

Minimization of $\mathcal{J}_{\mathrm{MR}\,\langle k+1\rangle}$ in (14) subject to (15) implies the following conditions on the Hamiltonian (16): $0 = \partial\mathcal{H}/\partial\Delta w^l_{\langle k\rangle}(t)$, $\lambda_{\langle k\rangle}(t) = \partial\mathcal{H}/\partial\Delta x_{\langle k\rangle}(t)$, and $\Delta x_{\langle k\rangle}(t+1) = \partial\mathcal{H}/\partial\lambda_{\langle k\rangle}(t+1)$. Differentiation and rearranging the former expression yields

$$\Delta w^l_{\langle k\rangle}(t) = \Gamma_1^{-1}\Big(\widetilde{D}^T\widetilde{W}_e\Big(\widetilde{e}(t) - \widetilde{C}\Delta x_{\langle k\rangle}(t)\Big) - W_w w^l_{\langle k\rangle}(t) - \widetilde{B}^T\lambda_{\langle k\rangle}(t+1)\Big). \quad (17)$$

Substitution of (17) into the latter two conditions yields the Hamiltonian system

$$\begin{bmatrix}\Delta x_{\langle k\rangle}(t+1)\\ \lambda_{\langle k\rangle}(t)\end{bmatrix} = \begin{bmatrix}H_{11} & H_{12}\\ H_{21} & H_{22}\end{bmatrix}\begin{bmatrix}\Delta x_{\langle k\rangle}(t)\\ \lambda_{\langle k\rangle}(t+1)\end{bmatrix} + \begin{bmatrix}E_{11} & E_{12}\\ E_{21} & E_{22}\end{bmatrix}\begin{bmatrix}\widetilde{e}_{\langle k\rangle}(t)\\ w^l_{\langle k\rangle}(t)\end{bmatrix} \quad (18)$$

with boundary conditions $\Delta x_{\langle k\rangle}(0) = 0$ and $\lambda_{\langle k\rangle}(N^l) = 0$. Using the well-known sweep method [17], the transformation $\lambda_{\langle k\rangle}(t) = P(t)\Delta x_{\langle k\rangle}(t) - g_{\langle k\rangle}(k)$ is used to remove the dependency of (18) on λ and to solve it for $\Delta x_{\langle k+1\rangle}(t)$. By observing that the resulting equation should hold for all $\Delta x_{\langle k\rangle}(t)$, it is clear that both (11) and (12) should hold. Next, using the same transformation in (17), subsequently substituting (15) and solving for $\Delta w^l_{\langle k\rangle}(t)$ yields

$$\Delta w^l_{\langle k\rangle}(t) = -L(t)\Delta x_{\langle k\rangle}(t) + L_e(t)\widetilde{e}(t) - L_w(t)w^l_{\langle k\rangle} + L_g(t)g_{\langle k\rangle}(t+1). \quad (19)$$

Next, the optimal state trajectory $\Delta x_{\langle k\rangle}$ can be computed by substituting (19) into (15) using boundary condition $\Delta x_{\langle k\rangle}(0) = 0$, leading to

$$\Delta x_{\langle k\rangle}(t+1) = (\widetilde{A} - \widetilde{B}L(t))\Delta x_{\langle k\rangle}(t) + \widetilde{B}L_g(t)g_{\langle k\rangle}(t+1) - \widetilde{B}L_w(t)w^l_{\langle k\rangle}(t) + \widetilde{B}L_e(t)\widetilde{e}(t). \quad (20)$$

Finally, (20), (19), and (13) yield the desired result (10). ∎

The main result of Theorem 9 is a state-space realization of an optimal ILC controller in the sense of (8). The order of the controller is equal to the state dimension of the underlying parametric model, see Theorem 5. The main computational effort in the implementation of the results of Theorem 9 are the computation of a discrete time Riccati equation (11) and backward recursion (12), whose dimensions depend on the state dimension of the model $\widetilde{\widetilde{J}}_{\mathrm{MR}}$, as well as the filtering by the time varying controller C^{ILC}, whose state dimension again equals the state dimension of the model $\widetilde{J}_{\mathrm{MR}}$. In this respect, increasing the trial length N^l, sampling ratio factor F, or the number of inputs n_u and outputs n_y of P only affects the length of the signals to be filtered. This enables an efficient implementation in a digital computer environment. In contrast, the approaches in [3], [4], [13], etc., resort to a linear least squares problem that involves matrices whose dimensions scale with N^l, F, n_u, and n_y.

Equation (14) reveals that the structure of the problem is related to the linear quadratic tracking problem [17], which is known to be not causal due to the need for future reference signals in the optimization. In the ILC problem, the error $\widetilde{e}_{\langle k\rangle}$ and command signal $w^l_{\langle k\rangle}$ from the previous trial constitute the reference signal, resulting in a causal control law in the trial domain. In contrast, the resulting ILC controller is generally not causal in the time domain. In analogy to the solution to the finite time linear quadratic optimal control problem [17], the controller C^{ILC} is time-varying. Alternatively, a time-invariant result can be derived, followed by an appropriate selection of boundary conditions, as is suggested in [18].

Although the results in Theorem 9 are related to the linear quadratic tracking problem [17] and the optimal ILC controllers in [14], [15], the derivation here is more involved due to the more complex criterion (9) and the presence of a direct feedthrough $\widetilde{D}$ in the multirate case, even if P is strictly proper, see Theorem 4.

The convergence properties of the ILC algorithm defined by Theorem 9 are similar to the results obtained in [3]. Hence, $w^l_{\langle 0\rangle}$ can be chosen arbitrarily due to monotonic convergence properties in w^l.

V. Example

A velocity feedback control controlled two mass-spring-damper system is considered with state-space realizations

$$P_o = \left[\begin{array}{ccc|c} -\frac{d}{m_1} & \frac{d}{m_1} & \frac{k}{m_1} & \frac{1}{m_1} \\ \frac{d}{m_2} & -\frac{d}{m_1} & -\frac{k}{m_2} & 0 \\ -1 & 1 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \end{array}\right], \qquad C^{d,l} = \left[\begin{array}{c|c} 0.6859 & 0.0132 \\ \hline 0.0132 & 10^{-4} \end{array}\right]$$

where $m_1 = m_2 = 4.8 \cdot 10^{-6}$, $k = 0.22$, and $d = 1 \cdot 10^{-4}$. In addition, $f^l = 500$ [Hz] and $f^h = 1500$ [Hz], hence $F = 3$.

Firstly, a model is identified using prediction error techniques, where the measured output is given by $e^h = -P^{d,h}\mathcal{H}_u(I+C^{d,l}P^{d,l})^{-1}w^l + v$, where v is a normally distributed zero mean white noise sequence. Then, Procedure 6 is applied and the 2-norm of the prediction error $\widetilde{\varepsilon}(t) = -\widetilde{e}(t) - \widetilde{\widetilde{J}}_{\mathrm{MR}}(q)w^l(t)$ is minimized, which is a standard open-loop LTI identification problem, see [16]. By virtue of Lemma 3 (a), $\|\widetilde{\varepsilon}(t)\|_2 = \|\mathcal{L}^{-1}\widetilde{\varepsilon}(t)\|_2$, hence the prediction error in the physical time domain is also minimized. By using independent validation data in the approach of [16, Section 16], it appears that the optimal model order for $\widetilde{\widetilde{J}}_{\mathrm{MR}}$ is four, which equals the order of the underlying closed-loop system.

Secondly, the optimal ILC procedure of Section IV is invoked. Here, $w^l_{\langle 0\rangle} = 0, W_e = 1, W_w = 0, W_\Delta = 10^{-6}$, and $N^l = 8$, hence $N^h = 24$. For comparison, a discrete time ILC procedure is considered, i.e., using $F = 1$. The results are depicted in Fig. 2. Here, $\bar{\Psi}(\omega)$ is the cumulative version of the normalized periodogram, which is defined as $\Psi(\omega) = (1/N)|X_N(\omega)|^2$ for a signal $x(t), t \in [0, N-1]$.

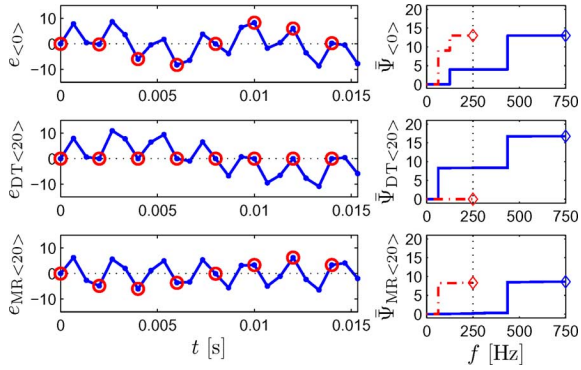

Fig. 2. Error signals. Left: time domain results with sampling frequencies f^h (dots) and f^l (circles). Right: normalized periodograms with sampling frequencies f^h (solid) and f^l (dash-dotted). Top: initial error. Middle: discrete time ILC. Bottom: proposed multirate ILC approach.

Firstly, observe that the initial error $e_{\langle 0\rangle}$ contains two dominant sinusoidal error components at 125 and 437.5 Hz, where the 437.5 Hz component appears as an aliased component at 62.5 Hz at the sampling frequency f^l, see Fig. 2, top right. Next, the discrete time ILC controller, which ignores the intersample response, results in a perfect response since $e^l_{\mathrm{DT}\langle 20\rangle}(t) = 0$. However, when considering the intersample response, it appears that the error $e^h_{\mathrm{DT}\langle 20\rangle}$ has increased compared to $e^h_{\langle 0\rangle}$. In fact, the criterion value $\mathcal{J}_{\mathrm{MR}}(w^l_{\langle 0\rangle}, e^h_{\langle 0\rangle}) = 624$ has increased to $\mathcal{J}_{\mathrm{MR}}(w^l_{\mathrm{DT}\langle 20\rangle}, e^h_{\mathrm{DT}\langle 20\rangle}) = 804$. Importantly, this deterioration cannot be observed from the at-sample response at f^l.

Finally, the proposed multirate approach achieves a performance $\mathcal{J}_{\mathrm{MR}}(w^l_{\mathrm{MR}\langle 20\rangle}, e^h_{\mathrm{MR}\langle 20\rangle}) = 413$, which is optimal. When considering the results in Fig. 2, bottom, it appears that the proposed multirate ILC controller results in a well-balanced at-sample and intersample response, both of which are significantly improved compared to $e_{\langle 0\rangle}$. Specifically, the multirate ILC controller does not attempt to attenuate the 437.5 Hz component, since this component originates from a higher frequency band than can be attenuated using w^l.

VI. Conclusion

In this technical note, system identification for ILC and an optimal ILC synthesis framework for sampled-data systems is presented. Time-varying aspects of the involved systems, which are introduced by the consideration of intersample behavior, are appropriately addressed in a multirate framework. The results address a largely unexplored area of identifying models in view of application in ILC. In addition, the simulation example reveals that the proposed multirate ILC approach outperforms common discrete time ILC algorithms. Specifically, it is shown that common discrete time ILC algorithms may decrease the performance of sampled-data systems by deteriorating the intersample response. In contrast, the proposed multirate approach optimally balances the at-sample and the intersample response. Although only a short trial length is considered in the simulation example, extensive additional simulations confirm that in contrast to prior results, the proposed ILC controller enables implementation for large-scale problems.

The proposed results pave the way for more efficient identification of models in view of ILC for sampled-data systems. Further research is still required to clarify the intimate connection between system identification and ILC, e.g., with respect to experiment design, the role of model uncertainty, and robust ILC synthesis.

REFERENCES

[1] T. Oomen, M. van de Wal, and O. Bosgra, "Design framework for high-performance optimal sampled-data control with application to a wafer stage," *Int. J. Control*, vol. 80, no. 6, pp. 919–934, 2007.

[2] R. W. Longman and C.-P. Lo, "Generalized holds, ripple attenuation, and tracking additional outputs in learning control," *J. Guid., Control Dyn.*, vol. 20, no. 6, pp. 1207–1214, 1997.

[3] T. Oomen, J. van de Wijdeven, and O. Bosgra, "Suppressing inter-sample behavior in iterative learning control," *Automatica*, vol. 45, no. 4, pp. 981–988, 2009.

[4] J. van de Wijdeven and O. Bosgra, "Hankel iterative learning control for residual vibration suppression with MIMO flexible structure experiments," in *Proc. 46th Conf. Dec. Control*, New Orleans, LA, 2007, pp. 258–263.

[5] H.-S. Ahn, K. L. Moore, and Y. Chen, "Monotonic convergent iterative learning controller design based on interval model conversion," *IEEE Trans. Autom. Control*, vol. 51, no. 2, pp. 366–371, Feb. 2006.

[6] M. Q. Phan and J. A. Frueh, "System identification and learning control," in *Iterative Learning Control: Analysis, Design, Integration and Applications*, Z. Bien and J. Xu, Eds. Norwell, MA: Kluwer Academic Publishers, 1998.

[7] R. W. Longman and M. Q. Phan, "Iterative learning control as a method of experiment design for improved system identification," *Opt. Meth. Softw.*, vol. 21, no. 6, pp. 919–941, 2006.

[8] H. Hjalmarsson, "From experiment design to closed-loop control," *Automatica*, vol. 41, pp. 393–438, 2005.

[9] F. Ding, L. Qiu, and T. Chen, "Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems," *Automatica*, vol. 45, no. 2, pp. 324–332, 2009.

[10] B. A. Bamieh and J. B. Pearson, "A general framework for linear periodic systems with applications to $\mathcal{H}_\infty$ sampled-data control," *IEEE Trans. Autom. Control*, vol. 37, no. 4, pp. 418–435, Apr. 1992.

[11] F. Ding and T. Chen, "Hierarchical identification of lifted state-space models for general dual-rate systems," *IEEE Trans. Circ. Syst. I*, vol. 52, no. 6, pp. 1179–1187, Jun. 2005.

[12] F. Ding, G. Liu, and X. P. Liu, "Partially coupled stochastic gradient identification methods for non-uniformly sampled systems," *IEEE Trans. Autom. Control*, vol. 55, no. 8, pp. 1976–1981, Aug. 2010.

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

[14] N. Amann, D. H. Owens, and E. Rogers, "Iterative learning control for discrete-time systems with exponential rate of convergence," *Proc. Inst. Elect. Eng., Control Theory Appl.*, vol. 143, no. 2, pp. 217–224, 1996.

[15] B. G. Dijkstra and O. H. Bosgra, "Extrapolation of optimal lifted system ILC solution, with application to a waferstage," in *Proc. Amer. Control Conf.*, Anchorage, AK, 2002, pp. 2595–2600.

[16] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1999.

[17] F. L. Lewis and V. L. Syrmos, *Optimal Control*, 2nd ed. New York: Wiley, 1995.

[18] E. Zattoni, "Structural invariant subspaces of singular Hamiltonian systems and nonrecursive solutions of finite-horizon optimal control problems," *IEEE Trans. Autom. Control*, vol. 53, no. 5, pp. 1279–1284, May 2008.

Stabilization of Switched Systems With Nonlinear Impulse Effects and Disturbances

Yuangong Sun

***Abstract*—In this note, we investigate the stabilization of switched systems with nonlinear impulse effects and disturbances. Based on the estimate on transition matrices and Gronwall inequality, feedback laws are designed to achieve the exponential stability of the closed-loop system with arbitrary switching frequency.**

***Index Terms*—Disturbance, impulse, stabilization, switched system.**

I. INTRODUCTION

Switched systems belong to a class of hybrid dynamic systems consisting of a family of continuous or discrete time subsystems and a switching law specifying the switches between them. In the last twenty years, the stability and stabilization problems for switched systems have attracted considerable efforts. When the switching laws are modeled as finite state Markov chains, necessary and sufficient conditions have been given to solve the problem for both the nonadaptive case (c.f. [1]) and the adaptive case (c.f. [2]). When the switching law is arbitrary, one way to investigate the stability and stabilization problems is to find a common Lyapunov function for all the subsystems (c.f. [3]–[6]). Another commonly used approach is to assume that a system remains unswitched for a period long enough to allow the overshoots of the closed-loop system in the transient phases to fade (c.f. [7] and [8]).

For systems that switch among a finite set of controllable linear systems, the stabilization problem of switched systems with arbitrary switching frequency has been studied in [9] and [10] by developing an improved estimation on transition matrices. For the case of planar switched systems, necessary and sufficient conditions for stabilization of the system under arbitrary switching law are given in [5] and [11]. When the system state is not easy to measure directly, the observer-based stabilization problem has been studied in [10] and [12]. For switched systems with impulse effects, exponential stability was analyzed by using the matrix measure concept and average dwell time approach in [13]. However, for switched systems with nonlinear impulse effects and disturbances, the stabilization problem of the switched system with arbitrary switching frequency remains unknown to the best of our knowledge.

In this note, we will follow the method formulated in [9] and [10] to further study the stabilization problem of switched systems with nonlinear impulse effects and disturbances. To guarantee the exponential stability of such a system at a given switching frequency, it is certainly not enough to just stabilize each individual system for the obvious reason that the overshoots due to impulses and disturbances may destroy the stability. Based on an improved estimate on

Manuscript received September 06, 2010; revised January 09, 2011, May 15, 2011; accepted June 08, 2011. Date of publication June 16, 2011; date of current version November 02, 2011. This work was supported by the National Natural Science Foundations of China (60704039, 61174217) and the Natural Science Foundations of Shandong Province (ZR2010AL002, JQ201119). Recommended by Associate Editor D. Liberzon.

The author is with School of Mathematics, Shandong Provincial Key Laboratory of Network Based Intelligent Computing, University of Jinan, Jinan 250022, China (e-mail: ss_sunyg@ujn.edu.cn).

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

Digital Object Identifier 10.1109/TAC.2011.2159889