Mechatronics 47 (2017) 134–147

Contents lists available at ScienceDirect

Mechatronics

journal homepage: www.elsevier.com/locate/mechatronics

Sparse iterative learning control with application to a wafer stage: Achieving performance, resource efficiency, and task flexibility☆

Tom Oomen[a,*], Cristian R. Rojas[b]

[a] Eindhoven University of Technology, Faculty of Mechanical Engineering, Control Systems Technology group, PO Box 513, Eindhoven 5600MB , The Netherlands
[b] Automatic Control Lab, Electrical Engineering, KTH – Royal Institute of Technology, Stockholm S-100 44, Sweden

ARTICLE INFO

Article history:
Received 5 June 2017
Revised 7 August 2017
Accepted 13 September 2017
Available online 6 October 2017

Keywords:
Iterative learning control
Motion control
Feedforward
Sparse optimization
Resource-efficient control

ABSTRACT

Trial-varying disturbances are a key concern in Iterative Learning Control (ILC) and may lead to inefficient and expensive implementations and severe performance deterioration. The aim of this paper is to develop a general framework for optimization-based ILC that allows for enforcing additional structure, including sparsity. The proposed method enforces sparsity in a generalized setting through convex relaxations using ℓ_1 norms. The proposed ILC framework is applied to the optimization of sampling sequences for resource efficient implementation, trial-varying disturbance attenuation, and basis function selection. The framework has a large potential in control applications such as mechatronics, as is confirmed through an application on a wafer stage.

1. Introduction

Iterative Learning Control (ILC) enables significant performance improvements for batch-to-batch control applications, by generating a command signal that compensates for repetitive disturbances through learning from previous iterations, also called batches or trials. Theoretical and implementation aspects, including convergence, causality, and robustness, have been addressed in, e.g., [1,12,41,44,45]. Furthermore, successful applications have been reported in, e.g., robotics [53], mechatronics [9], manufacturing [27], building control [43], nuclear fusion [19], and rehabilitation [20]. However, several disadvantages of present ILC frameworks that limit further applications include (i) high implementation cost due to highly unstructured command signals, which are expensive to implement; (ii) amplification of trial-varying disturbances, including measurement noise; (iii) inflexibility to changing reference trajectories. The aim of the present paper is to develop an ILC design framework that enforces sparsity, which enables addressing these aspects (i)–(iii).

Regarding (i) ILC typically generates signals that require a large number of command signal updates thus leading to an expensive implementation. ILC directly learns from measured signals that are contaminated by trial-varying disturbances such as measurement noise. These trial-varying disturbances are often modeled as a realization of a stochastic process [32]. As a result, the ILC command signals have infinite support. In sharp contrast, command signals that are obtained through traditional feedforward designs, including [31], have finite support and are highly sparse. Command signals with a high number of non-zero elements, or another appropriate structural constraint, may lead to a prohibitively expensive implementation, e.g., in wireless sensor networks, wireless control applications, or embedded platforms with shared resources [22]. Note that this is a different aspect than the actual computation of the command signal itself, which can be done in between subsequent tasks, see [60] for results in this direction.

Regarding (ii), ILC typically amplifies trial-varying disturbances. In fact, typical ILC approaches amplify these disturbances by a factor of two, as is shown in the present paper. Approaches to attenuate trial-varying disturbances include norm-optimal ILC with appropriate input weighting [12], higher-order ILC for addressing disturbances with trial-domain dynamics [24], and stochastic approximation-based ILC [14]. Also, a wavelet filtering-based approach is presented in [33], where a certain noise attenuation is achieved by setting certain wavelet coefficients to zero. In the present paper, a different approach is pursued to attenuate disturbances, where also wavelets immediately fit into the formulation, yet the sparsity can be enforced in an optimal way.

Regarding (iii), changing reference signals typically lead to performance degradation of ILC algorithms [5], since these essentially

☆ This paper was recommended for publication by Associate Editor Dr Peter Hehenberger.

* Corresponding author.
E-mail address: t.a.e.oomen@tue.nl (T. Oomen).

https://doi.org/10.1016/j.mechatronics.2017.09.004

constitute trial-varying disturbances. This is in sharp contrast to traditional feedforward designs [31] and is widely recognized in ILC designs. A basis task approach is proposed in [26], where the command input is segmented. A basis function framework is developed and applied in [8,34,54] using polynomial basis functions, which is further extended to rational basis functions in [61]. These basis functions are typically selected based on prior information, e.g., based on the approach in [31], and trial-and-error.

In model estimation and signal processing, the use of measured signals has comparable consequences, which has led to new regularization-based approaches that enforce sparsity. Early approaches include the non-negative garrote [10] and Least Absolute Shrinkage and Selection Operator (LASSO) [50]. These are further generalized in [3,13,25,52]. Related applications in system identification include [38,46].

Although important developments have been made in ILC and several successful applications have been reported, present approaches do not yet exploit the potential of enforcing additional structure and sparsity. The aim of the present paper is to develop a unified optimization-based approach to ILC that allows for explicitly enforcing structure and sparsity, enabling improved resource efficiency, disturbance attenuation, and flexibility to varying reference signals. The approach employs convex relaxations, enabling the use of standard optimization routines.

The main contribution of the present paper is a unified framework to sparse ILC. As subcontributions, trial-varying disturbances are analyzed in detail for explicit ILC algorithms (Section 3). Subsequently, a general optimization-based framework to sparse ILC is developed (Section 4), including many specific cases that are relevant to ILC applications. The results are confirmed through an application to a wafer stage system (Section 5). Related developments to the results in the present paper include the use of sparsity in control, where the main results have been related to Model Predictive Control (MPC), see [2,21,29].

Notation: Throughout, $\|x\|_{\ell_p}$ denotes the usual ℓ_p norm, $p \in \mathbb{Z}_{>0}$. Also, $\|x\|_0 = \sum_i \mathbf{1}(x_i \neq 0)$, i.e., the cardinality of x. Note that $\|x\|_0$ is not a norm, since it does not satisfy the homogeneity property. It relates to the general ℓ_p norm by considering the limit $p \to 0$ of $\|x\|_p$. A signal $x \in \mathbb{R}^N$ is called sparse if many of its components are zero, in which case $\|x\|_0 \ll N$. In addition, $\|\tilde{X}\|_{\mathcal{L}_\infty}$ and $\|\tilde{X}\|_{\mathcal{H}_\infty}$ denote the usual $\mathcal{L}_\infty$ and $\mathcal{H}_\infty$ norms of discrete time systems, respectively. Throughout, J denotes a system that maps an input space to an output space, operating either over finite or infinite time, which follows from the context. In certain cases, the system is assumed linear, time invariant, and scalar, with transfer function representation $\tilde{J}$. The power spectrum of a signal x is denoted ϕ_x, and is defined as in [32, Section 2.3].

2. Problem formulation

Consider the ILC system

$$e_j = r - Jf_j - v_j \tag{1}$$

be given, where $e_j \in \ell_2$ denotes the error signal to be minimized, $r \in \ell_2$ is the reference signal, $f_j \in \ell_2$ denotes the command signal, and $v_j \in \ell_2$ represents trial-varying disturbances, including measurement noise. Here and in the sequel, all signals are tacitly assumed to have appropriate dimensions. Furthermore, J represents the true system, either open-loop or closed-loop, with causal and stable transfer function $\tilde{J} \in \mathcal{RH}_\infty$. The index $j \in \mathbb{Z}_{\geq 0}$ refers to the trial number. Throughout, the command signal f_{j+1} is generated by an ILC algorithm

$$f_{j+1} = F(f_j, e_j), \tag{2}$$

where the ILC update F is defined in more detail later on. The general setup (1) encompasses the parallel ILC setup in Fig. 1, where

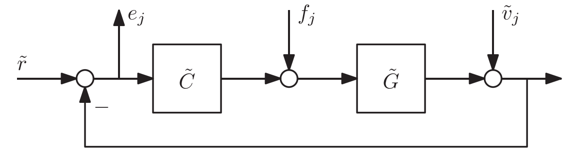

Fig. 1. Parallel ILC structure (3) as an example of (1).

$$e_j = S\tilde{r} - SGf_j - S\tilde{v}_j \tag{3}$$

where S follows from its transfer function $\tilde{S} = \frac{1}{1+\tilde{G}\tilde{C}}$, $r = S\tilde{r}$, $J = SG$, $v_j = S\tilde{v}_j$, and $\tilde{C}$, $\tilde{G}$ are assumed to be linear.

From (2) and (1), it is immediate that the trial-varying disturbance v_j directly affects the ILC command signal. In view of this observation, the problem investigated in this paper is to develop an ILC algorithm (2) that satisfies the following requirements:

R1) the iteration (1)-(2) is convergent over j;
R2) the iteration (1)-(2) leads to a small error e_j in the presence of trial-invariant disturbances r and trial-variant disturbances v_j;
R3) the resulting command signal f_j has a certain structure, including
  (a) a small $\|f_j\|_0$, and/or,
  (b) a piecewise constant f_j with a small number of jumps.

Here, R1 is a basic requirement for any ILC algorithm and ensures stability in the trial domain, in addition to the assumed stability in the time domain that is guaranteed by stability of J in (1), see also [45] for the stability of such two-dimensional systems. Requirement R2 essentially states that the ILC algorithm should effectively compensate for r, while avoiding amplification of trial-varying disturbances v_j. Requirement R3 is imposed to enable resource-efficient implementations in terms of sampling or communication requirements, depending on the particular application requirements.

3. Analysis of trial-varying disturbances in explicit ILC

The main contribution of this paper is a design framework for ILC that allows for enforcing sparsity and structure in the command signals, which in turn allows for explicitly addressing aspects (i)–(iii) as mentioned in Section 1. The main mechanism behind these aspects (i)–(iii) are trial-varying disturbances. First, ILC uses measured data, where trial-varying disturbances lead to highly unstructured command signals, which in turn are expensive to implement. Second, these trial-varying disturbances are amplified by typical ILC algorithms. Third, typical ILC algorithms are inflexible to changing reference signals, which can in fact be interpreted as trial-varying disturbances. Since trial-varying disturbances have a central role in all these aspects (i)–(iii), these trial-varying references are analyzed in typical ILC algorithms in this section.

In particular, explicit linear ILC algorithms of the general form

$$f_{j+1} = Q(f_j + Le_j) \tag{4}$$

are considered. The infinite time scalar case is considered, where Q: $\ell_2 \mapsto \ell_2$ and L: $\ell_2 \mapsto \ell_2$. Here, Q and L have associated transfer functions $\tilde{Q} \in \mathcal{RL}_\infty$ and $\tilde{L} \in \mathcal{RL}_\infty$. Note that $\tilde{J} \in \mathcal{RH}_\infty$ reflects causality and stability of the system. The fact that $\tilde{Q} \in \mathcal{RL}_\infty$ and $\tilde{L} \in \mathcal{RL}_\infty$ reflects that typical ILC algorithms are typically non-causal, and are usually implemented such that bounded solutions are obtained through finite-time preview or via stable inversion through a bilateral Z-transform [60].

The trial-varying disturbance v_j in (1) will propagate throughout the iterations through the iteration-domain update (4). The following assumption is widely adopted [32].