2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

Optimal Iterative Learning Control Design with Trial-varying Initial Conditions

Tong Duy Son, Goele Pipeleers and Jan Swevers

***Abstract*— In this paper we present an approach to deal with trial-varying initial conditions in norm-optimal iterative learning control (ILC). Varying initial conditions generally degrade the performance of conventional learning algorithms. We therefore introduce a worst-case optimization problem that accounts for trial-varying of initial conditions. The optimization is then reformulated as a convex minimization problem, which can be solved efficiently to generate the control signal. We investigate the relationship between the proposed approach and classical norm-optimal ILC; where we find that our methodology is equivalent to classical norm-optimal ILC with trial-varying parameters. Finally, simulation results of the presented technique are given.**

I. INTRODUCTION

Iterative learning control (ILC) has been widely adopted in control applications as an effective approach to improve the performance of repetitive processes [1]–[3]. The key idea of ILC is to update the control signal iteratively based on measured data from previous trials such that the output converges to the given reference trajectory. Most ILC algorithms assume identical, that is trial-invariant, initial conditions. In many applications however, perfect initial state resetting is not possible, which may degrade the performance of conventional ILC algorithms. This paper presents an ILC approach that is robust against trial-varying initial conditions.

The effects of non-repetitive initial conditions on the ILC controllers have been discussed in the literature. For example, in [4] it is shown that the P-type ILC performance error depends on the bound of the trial-varying initial state. An extension of this analysis is proposed in [5] for PD-type and PID-type ILC. [6] analyzes the stability of Lyapunov-based ILC for different initial resetting conditions, and [7] discusses the influence of variable initial state errors for average operator-based ILC. In addition, a robust ILC design using 2-D formulation for nonlinear systems with varying initial shift is considered in [8]. This work considers trial-varying initial conditions in a norm-optimal ILC [9] for linear time-invariant systems. The effect of random non-identical initializations on the performance of norm-optimal ILC is analyzed first. It is demonstrated that the tracking error depends on the initial states of both the current trial and all previous trials.

This research was supported by a Marie Curie Initial Training Network Grant within the 7th European Community Framework Programme. The authors of this work gratefully acknowledge support for this research under the project No. 264672 IMESCON provided by the EU.

T. D. Son, G. Pipeleers and J. Swevers are with the Department of Mechanical Engineering, Katholieke Universiteit Leuven, Celestijnenlaan 300B, B-3001 Heverlee, Belgium. Email: tong.duyson@mech.kuleuven.be.

Consequently, we propose a robust worst-case optimal ILC approach taking into account trial-varying initial conditions. The problem is formulated as a min-max problem with a quadratic cost function to minimize its worst-case value under trial-varying initial conditions. Even though there is already some research that considers a min-max formulation to design robust ILC taking into account model uncertainty [10]–[12], this work studies the varying initial conditions problem in ILC. Notice that [12] also discusses the problem of varying initial conditions, assuming that the difference between the initial state of two consecutive trials is a bounded random variable. Imposing this condition correlates the initial conditions between two consecutive trials and may cause on unbounded grow of the initial conditions in the trial domain. Furthermore, the developed worst-case ILC formulation can also be applied to the problem with trial-varying disturbances.

The remainder of this paper is organized as follows. Section II provides the background on norm-optimal ILC and analyzes the effect of trial-varying initial conditions on the performance of conventional norm-optimal ILC. Section III formulates the developed optimal ILC approach as a min-max problem, which minimizes the worst-case effect of trial-varying initial conditions. First an upper bound on this worst-case effect is found as the solution of a dual problem. Second a convex program is formulated to minimize this worst-case performance. Finally, as an additional contribution of the paper, the equivalence between the solution of the worst-case ILC algorithm and classical norm-optimal ILC with trial-varying weights is discussed in section IV. Simulation results are given in Section V, and Section VI concludes this paper.

II. PROBLEM FORMULATION

This section first describes the conventional norm-optimal ILC technique and subsequently discusses the influence of trial-varying initial conditions.

A. Norm-optimal ILC

Let the plant to be controlled be described by the following linear time invariant multi-input multi-output (MIMO) discrete-time state-space model:

$$\begin{aligned} x_j(k+1) &= Ax_j(k) + Bu_j(k) \\ y_j(k) &= Cx_j(k) \end{aligned} \tag{1}$$

where j is the trial index, N is the number of samples in a trial, and k is the time index, $k \in [0, N-1]$. Moreover, $x_j(k) \in \mathbb{R}^n$, $u_j(k) \in \mathbb{R}^m$, and $y_j(k) \in \mathbb{R}^p$ denote the state, control signal, and measured output respectively. In norm-optimal ILC, the system dynamics are generally reformulated in the lifted system framework by considering a finite time interval system representation as

$$\mathbf{y}_j = \mathbf{P}\mathbf{u}_j, \tag{2}$$

where input signal, output, and reference trajectory are defined as follows

$$\begin{aligned} \mathbf{u}_j &= [\ u_j^T(0) \quad u_j^T(1) \quad \dots \quad u_j^T(N-1)\ ]^T \\ \mathbf{y}_j &= [\ y_j^T(1) \quad y_j^T(2) \quad \dots \quad y_j^T(N)\ ]^T \\ \mathbf{y}_d &= [\ y_d^T(1) \quad y_d^T(2) \quad \dots \quad y_d^T(N)\ ]^T \end{aligned}$$

and the system model $\mathbf{P}$ is a lower triangular block Toeplitz matrix:

$$\mathbf{P} = \begin{bmatrix} p_1 & 0 & \cdots & 0 \\ p_2 & p_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ p_N & p_{N-1} & \cdots & p_1 \end{bmatrix}$$

with $p_i = CA^{i-1}B$. Here, we assume that the system has relative degree 1, i.e. $CB \neq 0$. Norm-optimal ILC is an optimization-based ILC design, where the control signal is computed by minimizing the following performance index:

$$\hat{J}(\mathbf{u}_{j+1}) = \|\mathbf{e}_{j+1}\|^2_{\mathbf{Q}} + \left\|\mathbf{u}_{j+1}\right\|^2_{\mathbf{S}} + \left\|\mathbf{u}_{j+1} - \mathbf{u}_j\right\|^2_{\mathbf{R}}, \tag{3}$$

where $\mathbf{Q}$, $\mathbf{R}$ are symmetric positive definite matrices, $\mathbf{S}$ is a symmetric positive semi-definite matrix. Moreover, we define $\|x\|^2 = x^T x$ and $\|x\|^2_{\mathrm{M}} = x^T \mathrm{M} x$. Note that $\hat{J}(\mathbf{u}_{j+1})$ is the nominal cost function without considering trial-varying initial conditions. In the cost function, $\mathbf{e}_{j+1}$ is the $(j+1)$-th trial tracking error, and is given by

$$\mathbf{e}_{j+1} = \mathbf{e}_j - \mathbf{P}(\mathbf{u}_{j+1} - \mathbf{u}_j). \tag{4}$$

Minimizing the cost (3) yields the following ILC update law:

$$\mathbf{u}_{j+1} = \mathbf{L}_u \mathbf{u}_j + \mathbf{L}_e \mathbf{e}_j \tag{5}$$

where

$$\begin{aligned} \mathbf{L}_u &= \left(\mathbf{P}^T\mathbf{Q}\mathbf{P} + \mathbf{S} + \mathbf{R}\right)^{-1}\left(\mathbf{P}^T\mathbf{Q}\mathbf{P} + \mathbf{R}\right) && (6) \\ \mathbf{L}_e &= \left(\mathbf{P}^T\mathbf{Q}\mathbf{P} + \mathbf{S} + \mathbf{R}\right)^{-1}\mathbf{P}^T\mathbf{Q}. && (7) \end{aligned}$$

ILC update law is asymptotically stable if

$$\rho(\mathbf{L}_u - \mathbf{L}_e\mathbf{P}) < 1, \tag{8}$$

where $\rho(A)$ is the spectral radius of matrix A. Furthermore, norm-optimal ILC is monotonically convergent if

$$\bar{\sigma}(\mathbf{L}_u - \mathbf{L}_e\mathbf{P}) < 1, \tag{9}$$

where $\bar{\sigma}(A)$ is the largest singular value of matrix A.

B. Effects of Trial-varying Initializations

The conventional ILC algorithms described in the previous section assumes trial-invariant system initialization. In practice, this is often difficult to realize, especially in case of systems with low damping or nonlinear friction. In order to analyze the effect of trial-varying system initialization, the system description (3) has to be extended to:

$$\mathbf{y}_j = \mathbf{P}\mathbf{u}_j + \Phi x_{0,j}, \tag{10}$$

where Φ denotes the observability matrix:

$$\Phi = \left[\ (CA)^T \quad (CA^2)^T \quad \dots \quad (CA^N)^T\ \right]^T, \tag{11}$$

and $x_{0,j}$ is the initial state at trial j. Here, $x_{0,j} \in \mathbb{R}^p$, hence independent of the trial length N. Without loss of generality, $\Phi x_{0,j}$ can be represented by $\mathbf{G}\mathbf{z}_j$, where the weight matrix $\mathbf{G}$ is an orthonormal matrix, that is $\mathbf{G}^T\mathbf{G} = \mathbf{I}$, and $\mathbf{z}_j$ is the trial-varying vector. This can be done using normalization, for instance, $\mathbf{G} = \Phi(\Phi^T\Phi)^{-0.5}$ and $\mathbf{z}_j = (\Phi^T\Phi)^{0.5}x_{0,j}$. The purpose of this change of matrices and vectors is to simplify the algorithm later on. Then (10) is reformulated as

$$\mathbf{y}_j = \mathbf{P}\mathbf{u}_j + \mathbf{G}\mathbf{z}_j. \tag{12}$$

Here, we assume that $\mathbf{z}_j$ is bounded by $\|\mathbf{z}_j\| \leq \alpha$. It is worth noting that this formulation can also be used for the case of trial-varying disturbances. Indeed, we can consider $\mathbf{z}_j$ as the trial-varying disturbance vector, while $\mathbf{G}$ is the filter matrix modeling the influence of the disturbances on the output. Therefore, trial-varying disturbances can be handled in a similar manner as trial-varying initial conditions.

Next, we examine the consequences of trial-varying initial conditions on the conventional norm-optimal ILC. Let us consider the classical norm-optimal ILC update law (5) for the system (12):

$$\begin{aligned} \mathbf{u}_{j+1} &= \mathbf{L}_u\mathbf{u}_j + \mathbf{L}_e\left(\mathbf{y}_d - \mathbf{P}\mathbf{u}_j - \mathbf{G}\mathbf{z}_j\right) \\ &= \left(\mathbf{L}_u - \mathbf{L}_e\mathbf{P}\right)\mathbf{u}_j - \mathbf{L}_e\mathbf{G}\mathbf{z}_j + \mathbf{L}_e\mathbf{y}_d, \end{aligned} \tag{13}$$

where $\mathbf{L}_u$ and $\mathbf{L}_e$ are chosen such that (9) is satisfied. The error at the $(j+1)$-th trial is now given as follows:

$$\begin{aligned} \mathbf{e}_{j+1} &= \mathbf{y}_d - \mathbf{P}\mathbf{u}_{j+1} - \mathbf{G}\mathbf{z}_{j+1} \\ &= \mathbf{y}_d - \mathbf{P}(\mathbf{L}_u\mathbf{u}_j + \mathbf{L}_e\mathbf{e}_j) - \mathbf{G}\mathbf{z}_{j+1}, \end{aligned} \tag{14}$$

or equivalently,

$$\begin{aligned} \mathbf{e}_{j+1} = (\mathbf{I} - \mathbf{P}\mathbf{L}_e)\mathbf{y}_d - \mathbf{P}(\mathbf{L}_u - \mathbf{L}_e\mathbf{P})\mathbf{u}_j \\ + \mathbf{P}\mathbf{L}_e\mathbf{G}\mathbf{z}_j - \mathbf{G}\mathbf{z}_{j+1}. \end{aligned} \tag{15}$$

The control signal in the first trial is set to zero without loss of generality. Then, by repeating the update equation (15), the $(j+1)$-th trial tracking error is expressed as a function of the initial conditions at the $(j+1)$-th trial and previous trials as follows:

$$\begin{aligned} \mathbf{e}_{j+1} &= (\mathbf{I} - \mathbf{P}\mathbf{L}_e)\mathbf{y}_d + \mathbf{P}\mathbf{L}_e\mathbf{G}\mathbf{z}_j - \mathbf{G}\mathbf{z}_{j+1} \\ &- \sum_{m=1}^{j-1}\mathbf{P}(\mathbf{L}_u - \mathbf{L}_e\mathbf{P})^{j-m}\mathbf{L}_e\mathbf{y}_d \\ &+ \sum_{m=1}^{j-1}\mathbf{P}(\mathbf{L}_u - \mathbf{L}_e\mathbf{P})^{j-m}\mathbf{L}_e\mathbf{G}\mathbf{z}_m. \end{aligned} \tag{16}$$

As can be seen from Eq. (16), the effects of varying initializations on the tracking error $\mathbf{e}_{j+1}$ are determined by the initial conditions at all trials $\mathbf{z}_{j+1}, \mathbf{z}_j, ..., \mathbf{z}_0$. Specifically, the initial conditions at the j-th and $(j+1)$-th trial have weightings $\mathbf{PL}_e\mathbf{G}$ and $-\mathbf{G}$ respectively. Hence, the contribution of $\mathbf{z}_{j+1}$ and $\mathbf{z}_j$ on the tracking error $\mathbf{e}_{j+1}$ is larger if they have opposite sign and $\mathbf{PL}_e$ is positive definite. Besides, the second summation term includes the influence of initial states from previous trials. Since $\|\mathbf{L}_u - \mathbf{L}_e\mathbf{P}\| \leq 1$, the older the trial, the less influence the initial state has on the error $\mathbf{e}_{j+1}$.

Accordingly, trial-varying initial conditions may degrade the performance of a learning controller. These reasons motivate us to design an ILC framework that can deal with trial-varying initial conditions.

III. Robust norm-optimal ILC algorithm

This Section discusses the design of a robust worst-case norm-optimal ILC approach against trial-varying initial conditions.

A. Formulation

Consider the cost function that accounts for trial-varying initial states:

$$J(\mathbf{u}_{j+1}, \mathbf{z}_{j+1}, \mathbf{z}_j) = \|\mathbf{e}_{j+1}\|^2_{\mathbf{Q}} + \left\|\mathbf{u}_{j+1}\right\|^2_{\mathbf{S}} + \left\|\mathbf{u}_{j+1} - \mathbf{u}_j\right\|^2_{\mathbf{R}}, \quad (17)$$

where the tracking error at trial $(j+1)$-th now includes the effect of trial-varying initial conditions:

$$\mathbf{e}_{j+1} = \mathbf{e}_j - \mathbf{P}(\mathbf{u}_{j+1} - \mathbf{u}_j) - \mathbf{G}(\mathbf{z}_{j+1} - \mathbf{z}_j). \quad (18)$$

We assume that the initial condition at each trial is a bounded random vector $\|\mathbf{z}_j\| \leq \alpha$, and uncorrelated in the trial domain. The optimization problem that is considered in the developed worst-case norm-optimal approach is the following problem:

$$\begin{array}{lll} \underset{\mathbf{u}_{j+1}}{\text{minimize}} & \underset{\mathbf{z}_{j+1}, \mathbf{z}_j}{\sup} & \{J(\mathbf{u}_{j+1}, \mathbf{z}_{j+1}, \mathbf{z}_j)\} \\ \text{subject to} & & \|\mathbf{z}_{j+1}\| \leq \alpha \\ & & \|\mathbf{z}_j\| \leq \alpha. \end{array} \quad (19)$$

Introducing a new variable $\mathbf{p}_{j+1}$ as

$$\mathbf{p}_{j+1} = \left[\begin{array}{cc} \mathbf{z}^T_{j+1} & \mathbf{z}^T_j \end{array}\right]^T, \quad (20)$$

we derive

$$\mathbf{e}_{j+1} = \mathbf{e}_j - \mathbf{P}(\mathbf{u}_{j+1} - \mathbf{u}_j) - \mathbf{H}\mathbf{p}_{j+1}, \quad (21)$$

where $\mathbf{J} = \left[\begin{array}{cc} \mathbf{I} & -\mathbf{I} \end{array}\right]$, and $\mathbf{H} = \mathbf{GJ}$. The cost function J turns out to be a function of the uncertain vector $\mathbf{p}_{j+1}$, where

$$\|\mathbf{p}_{j+1}\|^2 = \|\mathbf{z}_{j+1}\|^2 + \|\mathbf{z}_j\|^2 \leq 2\alpha^2. \quad (22)$$

Denote $\upsilon^2 = 2\alpha^2$, the control signal is consequently generated by solving the following optimization problem:

$$\begin{array}{lll} \underset{\mathbf{u}_{j+1}}{\text{minimize}} & \underset{\mathbf{p}_{j+1}}{\sup} & \{J(\mathbf{u}_{j+1}, \mathbf{p}_{j+1})\} \\ \text{subject to} & & \|\mathbf{p}_{j+1}\|^2 \leq \upsilon^2. \end{array} \quad (23)$$

Later, we will show that the two optimization problems (19) and (23) are equivalent since the worst-case $\mathbf{p}^*_{j+1}$ always satisfies the constraints in (19).

B. ILC algorithm

In order to find the worst-case cost function with respect to $\mathbf{p}_{j+1}$ in (23), let us consider the following maximization problem:

$$\begin{array}{ll} \underset{\mathbf{p}_{j+1}}{\text{maximize}} & \|\mathbf{e}_j - \mathbf{P}(\mathbf{u}_{j+1} - \mathbf{u}_j) - \mathbf{H}\mathbf{p}_{j+1}\|^2_{\mathbf{Q}} \\ \text{subject to} & \|\mathbf{p}_{j+1}\|^2 \leq \upsilon^2. \end{array} \quad (24)$$

Thanks to the S-procedure [13], strong duality holds for this maximization problem. Introducing the Lagrangian multiplier λ, the Lagrangian is expressed as

$$L(\mathbf{p}_{j+1}, \lambda) = \|\hat{\mathbf{e}}_{j+1} - \mathbf{H}\mathbf{p}_{j+1}\|^2_{\mathbf{Q}} + \lambda\left(\upsilon^2 - \|\mathbf{p}_{j+1}\|^2\right), \quad (25)$$

where $\hat{\mathbf{e}}_{j+1}$ denotes the nominal error:

$$\hat{\mathbf{e}}_{j+1} = \mathbf{e}_j - \mathbf{P}(\mathbf{u}_{j+1} - \mathbf{u}_j). \quad (26)$$

Maximization over $\mathbf{p}_{j+1}$ yields the following Lagrange dual function :

$$g(\lambda) = \left\{\begin{array}{ll} L(\mathbf{p}^*_{j+1}, \lambda) & \lambda\mathbf{I} - \mathbf{H}^T\mathbf{Q}\mathbf{H} \succeq 0, \\ & \mathbf{H}^T\mathbf{Q}\hat{\mathbf{e}}_{j+1} \in \mathcal{R}\left(\mathbf{H}^T\mathbf{Q}\mathbf{H} - \lambda\mathbf{I}\right) \\ +\infty & \text{otherwise} \end{array}\right. \quad (27)$$

where $\mathbf{I}$ is an identity matrix with suitable size, and

$$\mathbf{p}^*_{j+1} = \left(\mathbf{H}^T\mathbf{Q}\mathbf{H} - \lambda\mathbf{I}\right)^\dagger \mathbf{H}^T\mathbf{Q}\hat{\mathbf{e}}_{j+1}. \quad (28)$$

Applying the matrix inversion lemma with invertible C [14]: $(A+BCD)^{-1}BC = A^{-1}B(C^{-1}+DA^{-1}B)^{-1}$, we obtain

$$\left(\mathbf{H}^T\mathbf{Q}\mathbf{H} - \lambda\mathbf{I}\right)^\dagger \mathbf{H}^T\mathbf{Q} = \begin{bmatrix} -\mathbf{I} \\ \mathbf{I} \end{bmatrix} \mathbf{G}^T\left(\lambda\mathbf{Q}^{-1} - 2\mathbf{G}\mathbf{G}^T\right)^\dagger.$$

Hence, $\mathbf{p}^*_{j+1}$ generates $\mathbf{z}^*_{j+1} = -\mathbf{z}^*_j$, which implies $\left\|\mathbf{z}^*_j\right\| \leq \alpha$ and $\left\|\mathbf{z}^*_{j+1}\right\| \leq \alpha$ because of (22). In other words, the two optimization problems (19) and (23) are equivalent.

As a result, we obtain the Lagrange dual function as follows:

$$\begin{array}{rcl} g(\lambda) & = & \hat{\mathbf{e}}^T_{j+1}\mathbf{Q}\hat{\mathbf{e}}_{j+1} + \lambda\upsilon^2 \\ & + & \hat{\mathbf{e}}^T_{j+1}\mathbf{Q}\mathbf{H}\left(\lambda\mathbf{I} - \mathbf{H}^T\mathbf{Q}\mathbf{H}\right)^\dagger \mathbf{H}^T\mathbf{Q}\hat{\mathbf{e}}_{j+1}. \end{array} \quad (29)$$

The dual problem of (24) is then given by

$$\begin{array}{ll} \underset{\lambda}{\text{minimize}} & g(\lambda) \\ \text{subject to} & \lambda\mathbf{I} - \mathbf{H}^T\mathbf{Q}\mathbf{H} \succeq 0 \\ & \mathbf{H}^T\mathbf{Q}\hat{\mathbf{e}}_{j+1} \in \mathcal{R}\left(\mathbf{H}^T\mathbf{Q}\mathbf{H} - \lambda\mathbf{I}\right). \end{array} \quad (30)$$

Combining original minimization problem (23) with (30) yields

$$\begin{array}{ll} \underset{\mathbf{u}_{j+1}, \lambda}{\text{minimize}} & J_{\text{dual}}(\mathbf{u}_{j+1}, \lambda) \\ \text{subject to} & \lambda\mathbf{I} - \mathbf{H}^T\mathbf{Q}\mathbf{H} \succeq 0 \\ & \mathbf{H}^T\mathbf{Q}\hat{\mathbf{e}}_{j+1} \in \mathcal{R}\left(\mathbf{H}^T\mathbf{Q}\mathbf{H} - \lambda\mathbf{I}\right), \end{array} \quad (31)$$

where $J_{\text{dual}}(\mathbf{u}_{j+1}, \lambda)$ denotes the dual cost function,

$$J_{\text{dual}}(\mathbf{u}_{j+1}, \lambda) = \hat{\mathbf{e}}_{j+1}^T \mathbf{Q}\mathbf{H} \left(\lambda \mathbf{I} - \mathbf{H}^T\mathbf{Q}\mathbf{H}\right)^\dagger \mathbf{H}^T\mathbf{Q}\hat{\mathbf{e}}_{j+1} + \hat{\mathbf{e}}_{j+1}^T\mathbf{Q}\hat{\mathbf{e}}_{j+1} + \|\mathbf{u}_{j+1}\|_{\mathbf{S}}^2 + \left\|\mathbf{u}_{j+1} - \mathbf{u}_j\right\|_{\mathbf{R}}^2 + \lambda\upsilon^2. \quad (32)$$

J_{dual} is a convex function, hence (31) is a tractable optimization problem, and there exist efficient numerical algorithms to compute a globally optimal solution. Furthermore, the optimized input can be found from the equivalent optimization problem over an LMI using Schur complement and a slack variable t:

$$\begin{array}{ll} \underset{t, \mathbf{u}_{j+1}, \lambda}{\text{minimize}} & J_{\text{dual}}(\mathbf{u}_{j+1}, \lambda, t) \\ \text{subject to} & \begin{bmatrix} \lambda \mathbf{I} - \mathbf{H}^T\mathbf{Q}\mathbf{H} & \mathbf{H}^T\mathbf{Q}\hat{\mathbf{e}}_{j+1} \\ \hat{\mathbf{e}}_{j+1}^T\mathbf{Q}\mathbf{H} & t \end{bmatrix} \succeq 0, \end{array} \quad (33)$$

where

$$J_{\text{dual}}(\mathbf{u}_{j+1}, \lambda, t) = t + \hat{\mathbf{e}}_{j+1}^T\mathbf{Q}\hat{\mathbf{e}}_{j+1} + \|\mathbf{u}_{j+1}\|_{\mathbf{S}}^2 + \left\|\mathbf{u}_{j+1} - \mathbf{u}_j\right\|_{\mathbf{R}}^2 + \lambda\upsilon^2. \quad (34)$$

Note that the size of the LMI constraint depends on the number of the trial-varying states rather than on the number N of time samples in the trials. Moreover, input constraints can be imposed in the optimization problem, which is still a convex problem, i.e.,

$$\begin{array}{ll} \underset{t, \mathbf{u}_{j+1}, \lambda}{\text{minimize}} & J_{\text{dual}}(\mathbf{u}_{j+1}, \lambda, t) \\ \text{subject to} & \begin{bmatrix} \lambda \mathbf{I} - \mathbf{H}^T\mathbf{Q}\mathbf{H} & \mathbf{H}^T\mathbf{Q}\hat{\mathbf{e}}_{j+1} \\ \hat{\mathbf{e}}_{j+1}^T\mathbf{Q}\mathbf{H} & t \end{bmatrix} \succeq 0 \\ & |\mathbf{u}_{j+1}| \leq \overline{u} \\ & |\boldsymbol{\delta}\mathbf{u}_{j+1}| \leq \overline{\delta u}, \end{array} \quad (35)$$

where inequality constraints on $u_{j+1}(k)$ and $\delta u_{j+1}(k) = u_{j+1}(k) - u_{j+1}(k-1)$ are taken into account to avoid saturation of the actuators.

C. Special Case: $\mathbf{Q} = q\mathbf{I}$

This choice of weighting matrix is common in practice mainly because it simplifies the tuning of the norm-optimal ILC algorithm. An additional advantage of this choice is that it simplifies (32) yielding an analytic expression for λ. Combining the presented matrix inversion lemma, the condition $\mathbf{G}^T\mathbf{G} = \mathbf{I}$, and $\mathbf{Q} = q\mathbf{I}$, simplifies the dual cost function (32) to:

$$J_{\text{dual}}(\mathbf{u}_{j+1}, \lambda) = \hat{\mathbf{e}}_{j+1}^T \left(q\mathbf{I} + \frac{2q^2}{\lambda - 2q}\mathbf{G}\mathbf{G}^T\right)\hat{\mathbf{e}}_{j+1} + \|\mathbf{u}_{j+1}\|_{\mathbf{S}}^2 + \left\|\mathbf{u}_{j+1} - \mathbf{u}_j\right\|_{\mathbf{R}}^2 + \lambda\upsilon^2. \quad (36)$$

Minimizing this cost function with respect to λ yields:

$$\lambda = 2q + \frac{\sqrt{2}\left\|q\mathbf{G}^T\hat{\mathbf{e}}_{j+1}\right\|}{\upsilon}, \quad (37)$$

which satisfies the constraints in (27) since now $\left\|\mathbf{H}^T\mathbf{Q}\mathbf{H}\right\| = 2q$. Substituting Eq. (37) into (36), yields the following optimization problem with respect to $\mathbf{u}_{j+1}$:

$$\underset{\mathbf{u}_{j+1}}{\text{minimize}} \quad J_{\text{wc}}(\mathbf{u}_{j+1}) \quad (38)$$

where

$$J_{\text{wc}}(\mathbf{u}_{j+1}) = q\hat{\mathbf{e}}_{j+1}^T\hat{\mathbf{e}}_{j+1} + 2\sqrt{2}\upsilon\left\|q\mathbf{G}^T\hat{\mathbf{e}}_{j+1}\right\| + 2q\upsilon^2 + \|\mathbf{u}_{j+1}\|_{\mathbf{S}}^2 + \left\|\mathbf{u}_{j+1} - \mathbf{u}_j\right\|_{\mathbf{R}}^2. \quad (39)$$

The optimal control signal is found through minimizing $J_{\text{wc}}(\mathbf{u}_{j+1})$, which is a convex function.

IV. Robust vs classical norm-optimal ILC

This section discusses the relationship between the developed robust worst-case norm-optimal ILC approach and the classical norm-optimal ILC as described in section II.A. We first rewrite the problem (31) as

$$\begin{array}{ll} \underset{\lambda_{j+1}}{\text{minimize}}\ \underset{\mathbf{u}_{j+1}}{\text{minimize}} & J_{\text{dual}}(\mathbf{u}_{j+1}, \lambda_{j+1}) \\ \text{subject to} & \lambda_{j+1}\mathbf{I} - \mathbf{H}^T\mathbf{Q}\mathbf{H} \succeq 0 \\ & \mathbf{H}^T\mathbf{Q}\hat{\mathbf{e}}_{j+1} \in \mathcal{R}\left(\mathbf{H}^T\mathbf{Q}\mathbf{H} - \lambda_{j+1}\mathbf{I}\right). \end{array}$$

For the calculation of the optimal solution of this min-min problem, the optimal input is achieved by differentiating the cost function with respect to $\mathbf{u}_{j+1}$ yielding

$$\mathbf{u}_{j+1}^*(\lambda_{j+1}) = \left(\mathbf{P}^T\mathbf{Q}_{j+1}\mathbf{P} + \mathbf{R} + \mathbf{S}\right)^{-1}\mathbf{P}^T\mathbf{Q}_{j+1}\mathbf{e}_j + (\mathbf{P}^T\mathbf{Q}_{j+1}\mathbf{P} + \mathbf{R} + \mathbf{S})^{-1}(\mathbf{P}^T\mathbf{Q}_{j+1}\mathbf{P} + \mathbf{R})\mathbf{u}_j, \quad (40)$$

where $\mathbf{Q}_{j+1}$ is dependent on λ_{j+1}, and is given by

$$\mathbf{Q}_{j+1}(\lambda_{j+1}) = \mathbf{Q} + \mathbf{Q}\mathbf{H}\left(\lambda_{j+1}\mathbf{I} - \mathbf{H}^T\mathbf{Q}\mathbf{H}\right)^\dagger\mathbf{H}^T\mathbf{Q}. \quad (41)$$

After that, the optimal λ_{j+1}^* is found from the following optimization problem:

$$\begin{array}{ll} \underset{\lambda_{j+1}}{\text{minimize}} & J_{\text{dual}}\left(\mathbf{u}_{j+1}^*(\lambda_{j+1})\right) \\ \text{subject to} & \lambda_{j+1}\mathbf{I} - \mathbf{H}^T\mathbf{Q}\mathbf{H} \succeq 0 \\ & \mathbf{H}^T\mathbf{Q}\hat{\mathbf{e}}_{j+1} \in \mathcal{R}\left(\mathbf{H}^T\mathbf{Q}\mathbf{H} - \lambda_{j+1}\mathbf{I}\right). \end{array} \quad (42)$$

This is a nonlinear optimization problem with respect to λ_{j+1}. Once λ_{j+1}^* is calculated, the optimal learning gain $\mathbf{Q}_{j+1}(\lambda_{j+1})$ and the optimal input $\mathbf{u}_{j+1}^*(\lambda_{j+1})$ are obtained. In other words, these are a posteriori analyses. First, the dependency of $\mathbf{u}_{j+1}^*(\lambda_{j+1})$ on $\mathbf{Q}_{j+1}(\lambda_{j+1})$ is determined. After that, λ_{j+1}^* is generated from the optimization problem (42). Then $\mathbf{u}_{j+1}^*$ is updated with the given λ_{j+1}^*.

Comparing the robust algorithm (40) with classical norm-optimal ILC (5), shows that both use the same expressions to calculate the next trial's input, except for the weight $\mathbf{Q}_{j+1}$ which is updated trial-by-trail in the robust approach. Moreover, note that (36) and (37) implies

$$\mathbf{Q}_{j+1} = \mathbf{Q} + \frac{\sqrt{2}q\upsilon}{\|\mathbf{G}^T(\mathbf{e}_j - \mathbf{P}(\mathbf{u}_{j+1} - \mathbf{u}_j))\|}\mathbf{G}\mathbf{G}^T \quad (43)$$

if $\mathbf{Q} = q\mathbf{I}$. Hence, when the amount of uncertainty is very small, i.e. $\upsilon \approx 0$, the updated weight is approximately equal to the given $\mathbf{Q}$, and the larger the uncertainty, the larger $\|\mathbf{Q}_{j+1}\|$. The updated $\|\mathbf{Q}_{j+1}\|$ is larger than $\|\mathbf{Q}\|$ since the objective of worst-case norm-optimal ILC is to minimize the cost function with an additional part resulting from trial-varying initial conditions rather than only the nominal cost.

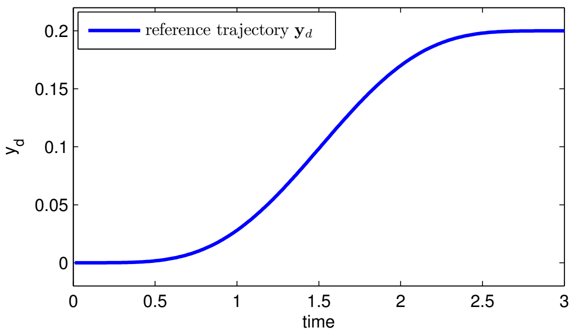

Fig. 1: Reference trajectory

V. Simulation

In this section, we simulate the proposed technique with a lowly damped second-order system:

$$P(s) = \frac{1}{s^2 + 0.02s + 1}. \quad (44)$$

The model is discretized with sampling time $T_s = 0.01s$, then lifted with $N = 300$ samples. The model has two trial-varying initial state variables $x_{01,j}$ and $x_{02,j}$, which are random but bounded, i.e. $\|x_{0,j}\| \leq \alpha$, where $\|x_{0,j}\|^2 = \|x_{01,j}\|^2 + \|x_{02,j}\|^2$. Here, we assume $\alpha = 0.005$. The control objective is to track a reference trajectory, which is a smoothed step function as shown in Fig. 1.

Simulations of the proposed worst-case ILC algorithm (31) are then performed. The weight matrices are simply selected as scaled identity matrices, $\mathbf{Q}, \mathbf{R}, \mathbf{S} = (1, 1, 0.1)\mathbf{I}$. In the first simulation, we analyze the effects of trial-varying initial states on the norm-optimal cost function. First, the initial states are assumed equal to zero while the control signal is a zero vector in the first trial without loss of generality. Next, we consider the cost function of the second trial as a function of the initial states x_{01} and x_{02} for both the proposed worst-case ILC and classical norm-optimal ILC, shown in Fig. 2.

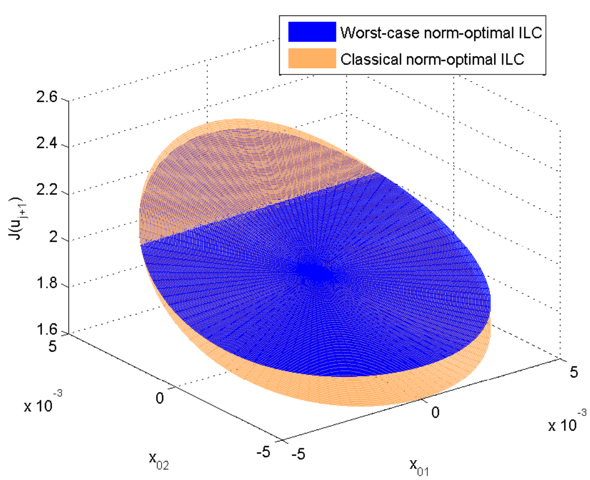

Fig. 2: Cost function with trial-varying initial states

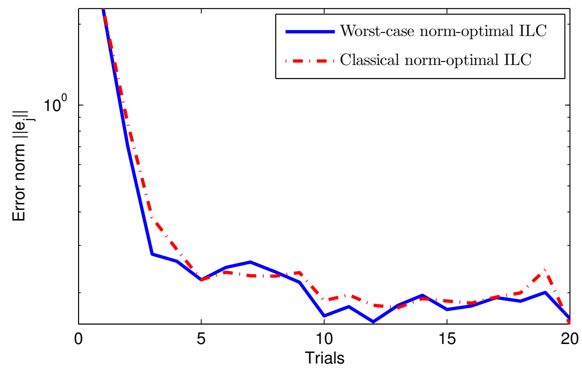

Fig. 3: Tracking error with trial-varying initial states

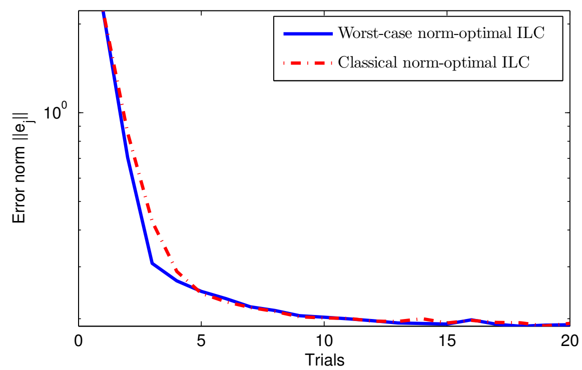

Fig. 4: Average performance error with 100 simulations

The results show that the robust worst-case ILC algorithm yields a smaller worst-case value than the classical norm-optimal ILC with respect to all bounded trial-varying initial states.

In order to validate the worst-case ILC in the trial-domain, trial-varying initial states are generated randomly within the given bound every trial. As a result, the tracking errors obtained with the worst-case optimal ILC controller for a random test case are shown in Fig. 3 (solid line). For comparison purposes, we also apply the same set of weight matrices $\mathbf{Q}, \mathbf{R}, \mathbf{S}$ and generated initial states to the classical norm-optimal ILC approach, and plot the results in Fig.3 (dashed line). Moreover, this simulation experiment is repeated 100 times and the average performance using both the worst-case ILC and classical norm-optimal ILC is plotted in Fig. 4. It can be seen that the proposed worst-case optimal ILC generally achieves faster convergence than the classical design.

Finally, we consider the equivalence between the robust design and norm-optimal ILC with trial-varying gains. Applying the equivalent adaptive ILC algorithm (40) for the given test case of Fig. 3, the trial-variation of $\|\mathbf{Q}_j\|$ is illustrated in Fig. 5. This figure confirms that $\|\mathbf{Q}_j\|$ is larger than $\|\mathbf{Q}\| = 1$ as discussed in Section IV.

VI. Conclusion

The concept of learning through experience of ILC has been extensively applied in the area of control. However,

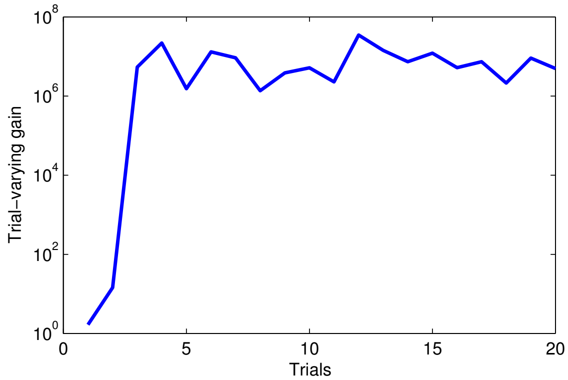

Fig. 5: Trial variation of $\|\mathbf{Q}_j\|$ for the test case in Fig. 3

when there are trial-varying initial states, these ILC approaches have trouble in generating satisfactory performance. In this paper, we have presented an analysis of the influence of trial-varying initial states on norm-optimal ILC system performance. After that, we have proposed a robust worst-case ILC algorithm, where the optimal input is computed efficiently via convex programming to deal with trial-varying initial conditions. It has been shown that the developed formulation can outperform existing optimal ILC in convergence speed. We also interpret the proposed controller through an equivalent adaptive norm-optimal ILC, yields trial-varying learning gains. The relation between robust ILC and adaptive norm-optimal ILC which respect to robustness and computation time will be studied in future work.

VII. Acknowledgments

This work was supported by the European Commission under the EU Framework 7 funded Marie Curie Initial Training Network (ITN) IMESCON (grant no. 264672). This work also benefits from project IWT-SBO 80032 (LECOPRO) of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen), project G.0377.09 of the Research Foundation - Flanders (FWO - Flanders), and from KU Leuven-BOF PFV/10/002 Cente of Excellence: Optimization in Engineering (OPTEC), and the Belgian Network Dynamical Systems, Control and Optimization (DYSCO), initiated by the Belgian Science Policy Office. Goele Pipeleers is Postdoctoral Fellow of the Research Foundation Flanders (FWO).

References

[1] H.S. Ahn, Y.Q. Chen, and K.L. Moore, "Iterative learning control: Brief survey and categorization", *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1099–1121, 2007.

[2] 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 Systems Magazine*, vol. 26, no. 3, pp. 96–114, 2006.

[3] K.L. Moore, "Iterative learning control: An expository overview", *Applied and Computational Control, Signals, and Circuits*, vol. 1, no. 1, pp. 151–214, 1999.

[4] G. Heinzinger, D. Fenwick, B. Paden, F. Miyazaki, "Stability of learning control with disturbances and uncertain initial conditions", *IEEE Transactions on Automatic Control*, vol. 37, no. 1, pp. 110–114, 1992.

[5] K. H. Park and Z. Bien, "A generalized iterative learning controller against initial state error", *International Journal of Control*, vol. 73, no. 10, pp. 871–881, 2000.

[6] J. Xu and R.Y. On, " Initial conditions in iterative learning control", *IEEE Transactions on Automatic Control*, vol. 50, no. 9, pp. 1349–1354, 2005.

[7] K. H. Park and Z. Bien, "A study on the effect of variable initial state error in average operator-based iterative learning control ", *in 17th IFAC World Congress*, 2008, vol. 17, no. 1, pp. 12448–12453.

[8] D. Meng, Y. Jia, J. Du, and S. Yuan, " Robust discrete-time iterative learning control for nonlinear systems with varying initial state shifts", *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2626–2631, 2009.

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

[10] D. H. Nguyen and D. Banjerdpongchai, "A convex optimization approach to robust iterative learning control for linear systems with time-varying parametric uncertainties," *Automatica*, vol. 47, no. 9, pp. 2039–2043, 2011.

[11] J. J. M. van de Wijdeven, M. C. F. Donkers and O. H. Bosgra, "Iterative learning control for uncertain systems: Noncausal finite time interval robust control design," *Int. J. Robust. Nonlinear Control*, vol. 21, no. 14, pp. 1645–1666, 2011.

[12] A. Haber, P.R. Fraanje, and M. Verhaegen, "Fast and robust iterative learning control for lifted systems", *in 18th IFAC World Congress*, 2011, vol. 18, no. 1, pp. 3617–3622.

[13] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge University Press, 2004.

[14] H. V. Henderson and S. R. Searle, "On Deriving the Inverse of a Sum of Matrices", *SIAM Review*, vol. 23, no. 1, pp. 53-60, Jan. 1981.