Nonlinear iterative learning control for discriminating between disturbances

Aarnoudse, Leontine; Pavlov, Alexey; Oomen, Tom

DOI
10.1016/j.automatica.2024.111902

Publication date
2025

Document Version
Final published version

Published in
Automatica

Citation (APA)
Aarnoudse, L., Pavlov, A., & Oomen, T. (2025). Nonlinear iterative learning control for discriminating between disturbances. *Automatica*, *171*, Article 111902. https://doi.org/10.1016/j.automatica.2024.111902

Automatica 171 (2025) 111902

Contents lists available at ScienceDirect

Automatica

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

Brief paper

Nonlinear iterative learning control for discriminating between disturbances ☆

Leontine Aarnoudse [a,*], Alexey Pavlov [b], Tom Oomen [a,c]

[a] Dept. of Mechanical Engineering, Control Systems Technology, Eindhoven University of Technology, Eindhoven, The Netherlands
[b] Dept. of Geoscience and Petroleum, Norwegian University of Science and Technology, Trondheim, Norway
[c] Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands

ARTICLE INFO

Article history:
Received 5 September 2023
Received in revised form 22 February 2024
Accepted 1 July 2024
Available online xxxx

Keywords:
Iterative learning control
Nonlinear control
Feedforward control
Stability of nonlinear systems

ABSTRACT

Disturbances in iterative learning control (ILC) may be amplified if these vary from one iteration to the next, and reducing this amplification typically reduces the convergence speed. The aim of this paper is to resolve this trade-off and achieve fast convergence, robustness and small converged errors in ILC. A nonlinear learning approach is presented that uses the difference in amplitude characteristics of repeating and varying disturbances to adapt the learning gain. Monotonic convergence of the nonlinear ILC algorithm is established, resulting in a systematic design procedure. Application of the proposed algorithm demonstrates both fast convergence and small errors.

1. Introduction

Iterative learning control (ILC) can attenuate repeating disturbances completely, yet it also amplifies iteration-varying disturbances. In ILC, the input that compensates the repeating disturbances is updated iteratively, resulting in high performance (Bristow, Tharayil, & Alleyne, 2006). To achieve this high performance, typical requirements for the ILC algorithm are the following:

R1. monotonic convergence (Longman, 2000),
R2. robustness against model uncertainty (Oomen, 2020),
R3. full attenuation of iteration-invariant disturbances,
R4. minimal amplification of iteration-varying disturbances (Butcher, Karimi, & Longchamp, 2008; Oomen & Rojas, 2017),
R5. fast convergence, and
R6. explicit ILC filters.

Thus, the aim of ILC is fast and robust monotonic convergence to small errors. These requirements lead to trade-offs and performance limitations. In particular, R3 and R4 are often conflicting, e.g., iteration-invariant noise is typically amplified up to a factor two (Butcher et al., 2008; Oomen & Rojas, 2017), in contrast to the iteration-invariant disturbances which are fully compensated.

In widely-used linear time-invariant (LTI) ILC approaches, limiting the amplification of iteration-varying disturbances typically results in reduced attenuation of iteration-invariant disturbances (Bristow et al., 2006) and slower convergence (Butcher et al., 2008). Approaches include using a low-pass robustness filter, systematically reducing an iteration-dependent learning gain (Butcher et al., 2008), or a combination such as a frequency-dependent learning gain. While robustness filters are common in ILC to ensure convergence (Bristow et al., 2006; Oomen, 2020), using this filter to limit noise amplification may also reduce the attenuation of repeating disturbances, which is undesired by R3. Reducing the learning gain reduces the amplification of the iteration-varying disturbances, but at the cost of slow convergence (R5). In norm-optimal ILC (Gunnarsson & Norrlöf, 2001), weights on the input signal and the change in input have effects similar to respectively robustness filters and reduced learning gains. For repetitive control, a related technique, Kalman filters may reduce the amplification of non-repeating disturbances (Longman, 2010), yet this depends strongly on accurate models of the system and the repeating disturbance.

In an attempt to circumvent the limitations of LTI ILC, adaptive ILC strategies aim to differentiate between iteration-varying and iteration-invariant disturbances based on their frequency or amplitude characteristics. In Oomen and Rojas (2017), the amplification of high-frequency iteration-varying disturbances is limited by adding a convex relaxation of the ℓ_0-norm of the input signal to the standard norm-optimal ILC criterion. While this sparse ILC algorithm leads to fast convergence, the input

☆ This work is part of the research programme VIDI with project number 15698, which is (partly) financed by the NWO, The Netherlands. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Daniele Astolfi under the direction of Editor Luca Zaccarian.

∗ Corresponding author.
E-mail addresses: l.i.m.aarnoudse@tue.nl (L. Aarnoudse), alexey.pavlov@ntnu.no (A. Pavlov), t.a.e.oomen@tue.nl (T. Oomen).

https://doi.org/10.1016/j.automatica.2024.111902

update that minimizes the criterion cannot be obtained in closed form, i.e., requirement R6 is not met, and as a result analyzing and enforcing robustness is challenging.

Nonlinear ILC designs using amplitude-dependent operators can exploit the typical amplitude characteristics of iteration-invariant and iteration-varying disturbances. This idea relates to the nonlinear filters used in variable-gain feedback controllers, see, e.g., Heertjes and Steinbuch (2004) or the performance analysis in Pavlov, Hunnekens, Wouw, and Nijmeijer (2013). In Heertjes and Tso (2007), a related idea is presented by including a deadzone in the ILC learning filter. In Heertjes, Rampadarath, and Waiboer (2009), a low-pass filter is included in the learning filter, but this implementation does not increase the robustness against model uncertainty (Bristow et al., 2006). The approach in Heertjes et al. (2009) does not allow for the use of a filter that does provide robustness (R2). Also, the convergence analysis cannot guarantee monotonic convergence (R1), such that excessive learning transients may occur (Longman, 2000). This limits the applicability in practice. A similar frequency-domain approach is introduced in Aarnoudse, Pavlov, and Oomen (2023), which assumes infinite-time signals and provides a conservative contraction-based monotonic convergence condition, which is difficult to meet in practice. Most analyses of nonlinear ILC consider ILC applied to nonlinear systems, e.g., Ahn, Choi, and Kim (1993) use an ILC update based on the relative degree of a nonlinear system, Volckaert, Diehl, and Swevers (2013) approximate a nonlinear system by a nominal model with an estimated correction term in norm-optimal ILC, and Altın and Barton (2017) study ILC for nonlinear systems as a two-dimensional differential repetitive process. In contrast, the current paper uses a nonlinear ILC algorithm to introduce an additional degree of design freedom for a linear system, a case to which these existing approaches cannot be applied directly.

Although major steps have been taken towards ILC algorithms that achieve fast convergence, strong attenuation of iteration-invariant disturbances and limited amplification of iteration-varying disturbances, an approach that also enables tuning for robustness and monotonic convergence is lacking. In this paper, a nonlinear lifted ILC algorithm is developed that includes a deadzone in the learning filter, resulting in fast convergence to small errors. The ILC algorithm enables tuning for robustness. Criteria for monotonic convergence follow from interpreting the ILC system as a MIMO discrete-time Lur'e system. The contribution is threefold.

- A nonlinear ILC algorithm is developed that addresses the requirements R1–R6.
- Full monotonic convergence proofs are provided, leading to criteria that are applied in a design framework.
- The approach is validated in simulation.

In addition, the propagation of iteration-varying and iteration-invariant disturbances in lifted ILC is analyzed, and the notion of monotonic convergence as commonly used in LTI ILC is extended to develop a convergence criterion for the nonlinear ILC algorithm. The extension is based on the notion of discrete-time nonlinear convergent systems (Pavlov & Van De Wouw, 2012), and enables the convergence analysis of the nonlinear ILC algorithm in the presence of iteration-varying disturbances.

This paper is structured as follows. In Section 2, the problem is introduced. In Section 3, the amplification of iteration-varying disturbances in ILC is analyzed. In Section 4, the nonlinear ILC algorithm is introduced and convergence conditions are developed. Section 5 demonstrates the selection of the design parameters. The approach is validated in simulation in Section 6 and conclusions are given in Section 7.

Notation: Throughout, $\|x\|_{\mathbf{P}} = \sqrt{x^{\top}\mathbf{P}x}$ for $\mathbf{P} \succ 0$ and $\mathbf{P} = \mathbf{I}$ leads to the 2-norm $\|x\|_2 = \sqrt{\sum_{i=1}^{N} |x_i|^2}$ for a vector $x \in \mathbb{R}^N$. The norm induced by the $\|x\|_{\mathbf{P}}$ vector norm for matrix $\mathbf{G} \in \mathbb{R}^{N\times N}$ is denoted by $\|\mathbf{G}\|_{\mathbf{P}}$, and $\mathbf{P} = \mathbf{I}$ gives the induced 2-norm $\|\mathbf{G}\|_{i2} = \bar{\sigma}(\mathbf{G})$. The vector ℓ_∞-norm is denoted by $\|x\|_{\ell_\infty} = \sup_n |x(n)|$. The spectral radius of a matrix is denoted by $\rho(\mathbf{G}) = \max_i |\lambda_i(\mathbf{G})|$, with $\lambda_i(\mathbf{G})$ the ith eigenvalue of $\mathbf{G}$. The sets of real, integer and non-negative integer numbers are denoted by $\mathbb{R}$, $\mathbb{Z}$ and $\mathbb{Z}_{\geq 0}$, respectively. The notation x_j, $j \in \mathbb{Z}_{\geq 0}$ indicates vector x in iteration j, and $x_j(k)$ denotes the kth sample in x_j. The lifted representation $\mathbf{J} \in \mathbb{R}^{N\times N}$ of a discrete-time LTI system J with Markov parameters h_i, $i = 0, 1, \ldots, N-1$ is given by

$$\mathbf{J} = \begin{bmatrix} h_0 & 0 & \ldots & 0 \\ h_1 & h_0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h_{N-1} & h_{N-2} & \ldots & h_0 \end{bmatrix}, \tag{1}$$

see, e.g., Gunnarsson and Norrlöf (2001) for an example.

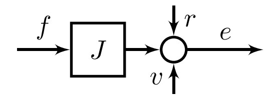

Fig. 1. Schematic representation the considered control scheme.

2. Problem formulation

Iterative learning control updates an input signal to learn to attenuate repeating disturbances. The error $e \in \mathbb{R}^N$ for the SISO discrete-time LTI system in Fig. 1 is given by

$$e = r - \mathbf{J}f - v, \tag{2}$$

with $\mathbf{J}$ according to (1), input signal $f \in \mathbb{R}^N$ and disturbances $r \in \mathbb{R}^N$ and $v \in \mathbb{R}^N$. The aim of ILC is to iteratively learn an input f that minimizes e in terms of the 2-norm $\|e\|_2$, leading to signals f_j and e_j that depend on iteration $j \in \mathbb{Z}_{\geq 0}$. Disturbance v_j is iteration-varying and disturbance r is constant over iterations. For example, r may be due to an iteration-invariant reference signal, and v_j may result from iteration-varying sensor noise. It is assumed that v_j is bounded, and that iteration-invariant and iteration-varying disturbances have distinct amplitude characteristics, i.e., $\|v_j\|_{\ell_\infty} \ll \|r\|_{\ell_\infty}\ \forall j$.

This paper aims to develop an ILC algorithm that minimizes $\|e\|_2$ in (2) in a small number of iterations. Iteration-invariant disturbance r should be attenuated completely without amplifying iteration-varying disturbance v_j, and the algorithm should ensure good learning transients, e.g., fast monotonic convergence, while being robust against model uncertainty.

To this end, the propagation of disturbances in ILC is analyzed in Section 3, illustrating that fast and complete compensation of r requires a high learning gain, while limiting the amplification of v_j requires a small learning gain. In Section 4, a nonlinear ILC algorithm is developed that differentiates between r and v_j based on their amplitude characteristics to apply time- and iteration-varying learning gains, and convergence conditions for this algorithm are provided.

3. Design trade-off in linear ILC

In this section, the propagation of iteration-varying and iteration-invariant disturbances in the ILC system (2) is analyzed. In addition, the trade-off between reduced amplification of disturbances and convergence speed is illustrated.

3.1. Lifted iterative learning control

Consider a linear ILC system of the form (2) for which the input signal f_j is updated iteratively according to

$$f_{j+1} = \mathbf{Q}(f_j + \alpha \mathbf{L} e_j). \tag{3}$$

Matrices $\mathbf{Q}$ and $\mathbf{L}$ follow from a norm-optimal ILC cost function that weights the 2-norms of the error and input signals (Gunnarsson & Norrlöf, 2001). The weighting on the change in input signal, which is typically included in this cost function, is replaced by a learning gain α which is chosen to be $\in (0, 1]$, analogous to other ILC methods (Bristow et al., 2006). This is elaborated upon in Section 5.

Eq. (2) can be substituted in (3) to obtain the following input and error iterations.

$$f_{j+1} = \mathbf{Q}(\mathbf{I} - \alpha \mathbf{L}\mathbf{J}) f_j + \alpha \mathbf{Q}\mathbf{L}(r - v_j), \tag{4}$$

$$e_{j+1} = (\mathbf{J}\mathbf{Q}\mathbf{J}^{-1} - \alpha \mathbf{J}\mathbf{Q}\mathbf{L}) e_j + (\mathbf{I} - \mathbf{J}\mathbf{Q}\mathbf{J}^{-1}) r + \mathbf{J}\mathbf{Q}\mathbf{J}^{-1} v_j - v_{j+1}. \tag{5}$$

This expression for the error iteration requires that $\mathbf{J}$ is invertible for the substitution $f_j = \mathbf{J}^{-1}(r - e_j - v_j)$ to be valid.

3.2. Analysis of disturbances in lifted ILC

Next, the influence of the design parameters $\mathbf{Q}$, $\mathbf{L}$ and α on the propagation of disturbances in linear ILC is analyzed in the stochastic setting. To this end, the following is assumed for iteration-varying disturbance v_j.

Assumption 1. The iteration-varying disturbance v_j is zero-mean, stationary and bounded. In addition, it is independent and identically distributed (i.i.d.) in the iteration domain.

Norm-optimal ILC aims to achieve a small error in terms of the 2-norm $\|e_j\|_2^2$, the expected value of which is given by

$$E\{\|e_j\|_2^2\} = \frac{1}{N}\sum_{k=1}^{N} E\{\tilde{e}_j^2(k)\} + E\{e_j(k)\}^2 \tag{6}$$

$$= E\{\tilde{e}_j^{\top}\tilde{e}_j\} + E\{e_j\}^{\top} E\{e_j\}, \tag{7}$$

since $E\{\|e_j\|_2^2\} = \frac{1}{N}\sum_{k=1}^{N} E\{e_j^2(k)\}$ by linearity and the variance can be expressed as

$$E\{\tilde{e}_j^2(k)\} = E\{e_j^2(k)\} - E\{e_j(k)\}^2. \tag{8}$$

Thus, a small 2-norm of the error signal requires that both the variance and expected value of the error are small and therefore they should both be taken into account in ILC. The following theorem shows how the variance and expected value of the error, which together determine the expected value of the error 2-norm according to (6), depend on $\mathbf{Q}$, $\mathbf{L}$ and α. A related statistical analysis using transfer functions with finite-time signals is developed in Butcher et al. (2008).

Theorem 2. *Consider the error iteration (5) with v_j according to Assumption 1, and assume that $\|\mathbf{J}\mathbf{Q}\mathbf{J}^{-1} - \alpha \mathbf{J}\mathbf{Q}\mathbf{L}\|_{\mathbf{P}} < \lambda$ for some $\mathbf{P} = \mathbf{P}^{\top} \succ 0$ and $\lambda \in (0, 1)$. Then, the expected value of the converged error is given by*

$$\lim_{j\to\infty} E\{e_j\} = (\mathbf{I} - \mathbf{J}\mathbf{Q}(\mathbf{J}^{-1} - \alpha \mathbf{L}))^{-1}(\mathbf{I} - \mathbf{J}\mathbf{Q}\mathbf{J}^{-1}) r. \tag{9}$$

In addition,

$$E\{\tilde{e}_{j+1}^{\top}\tilde{e}_{j+1}\} = E\{(\mathbf{J}\mathbf{Q}\mathbf{J}^{-1} v_j)^{\top}\mathbf{J}\mathbf{Q}\mathbf{J}^{-1} v_j\} + E\{v_{j+1}^{\top} v_{j+1}\} + E\{(\mathbf{J}\mathbf{Q}(\mathbf{J}^{-1} - \alpha \mathbf{L})\tilde{e}_j)^{\top}(\mathbf{J}\mathbf{Q}(\mathbf{J}^{-1} - \alpha \mathbf{L})\tilde{e}_j)\} - 2E\{(\mathbf{J}\mathbf{Q}\mathbf{J}^{-1} v_j)^{\top}\mathbf{J}\mathbf{Q}(\mathbf{J}^{-1} - \alpha \mathbf{L}) v_j\}. \tag{10}$$

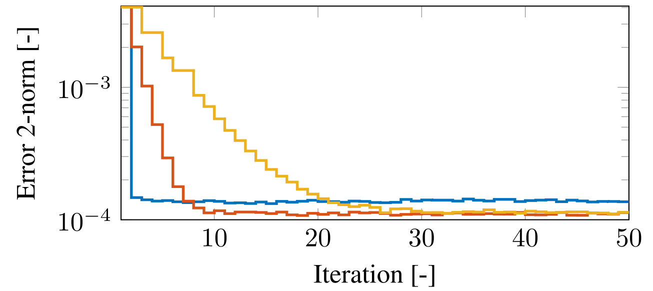

Fig. 2. Error 2-norm over iterations for $\alpha = 1$ (—), 0.5 (—) and 0.2 (—), averaged over 20 realizations. Small learning gains lead to smaller converged errors at the cost of slower convergence.

The proof of Theorem 2 is given in the Appendix. In the case that $\mathbf{L} = \mathbf{J}^{-1}$ and $\mathbf{Q} = \mathbf{I}$, the variance of the converged error at sample k is given by

$$E\{\tilde{e}(k)_{j+1}^2\} = E\{v_j(k)^2\} + E\{v_{j+1}(k)^2\} + (1-\alpha)^2 E\{\tilde{e}_j(k)^2\} - 2(1-\alpha)E\{v_j(k)^2\}. \tag{11}$$

For $j \to \infty$, it holds that $E\{\tilde{e}_{j+1}(k)^2\} = E\{\tilde{e}_j(k)^2\}$. Using $E\{v_j(k)\} = E\{v_{j+1}(k)\} = \sigma_v^2$, the variance of the converged error is therefore given by

$$E\{(\tilde{e}_\infty(k))^2\} = \frac{2}{2-\alpha}\sigma_v^2. \tag{12}$$

Eq. (12) illustrates that the variance of the error can be reduced by reducing the learning gain α. In particular, for $\alpha \to 0$, $E\{(\tilde{e}(k)_\infty)^2\} \to \sigma_v^2$, which is the smallest variance that can be achieved, since f_j cannot compensate the unknown disturbance v_j in iteration j. However, reducing α when $\mathbf{Q} \neq \mathbf{I}$ may increase the expected value of the converged error as well, as illustrated by substituting $\mathbf{Q} \neq \mathbf{I}$ and $\mathbf{L} = \mathbf{J}^{-1}$ in (9). Moreover, reducing α reduces the convergence speed significantly. This trade-off between fast convergence and small converged errors is illustrated in Fig. 2, using simulations that are elaborated upon in Section 6.

4. Nonlinear ILC

In Section 3, a trade-off between amplification of iteration-varying disturbances, attenuation of iteration-invariant disturbances and convergence speed has been illustrated. In this section a nonlinear operator is implemented that alleviates this trade-off. First, nonlinear ILC is introduced. Second, the notion of exponential convergence for discrete-time (nonlinear) systems is introduced and related to the standard requirement of monotonic convergence in ILC. This notion is used to develop a convergence criterion for nonlinear ILC.

4.1. Nonlinear ILC

To resolve the trade-off in Section 3, a deadzone nonlinearity is included in the ILC update. The idea is that this nonlinearity leads to different learning gains depending on amplitude of the error signal. For example, small learning gains are applied to low-amplitude error signals, limiting the amplification of iteration-varying disturbances, and high learning gains are applied to high-amplitude error signals resulting in fast attenuation of large iteration-invariant disturbances. The deadzone nonlinearity is implemented as follows. Consider the system (2) with an input update of the form

$$f_{j+1} = \mathbf{Q}(f_j + \alpha \mathbf{L} e_j + \mathbf{L}\varphi(e_j)), \tag{13}$$

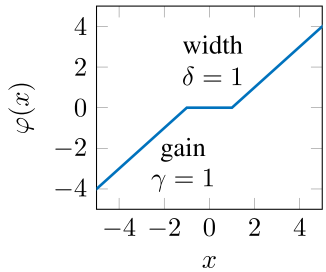

Fig. 3. Deadzone nonlinearity with $\gamma = 1$, $\delta = 1$.

with vector-valued deadzone nonlinearity $\boldsymbol{\varphi}(e_j)$, which applies a scalar deadzone $\varphi(e_j(k))$, illustrated in Fig. 3, to each entry $e_j(k)$, $k = 1, 2, \ldots, N$ of $e_j \in \mathbb{R}^N$ according to

$$\varphi(e_j(k)) = \begin{cases} 0, & \text{if } |e_j(k)| \le \delta \\ \left(\gamma - \frac{\gamma\delta}{|e_j(k)|}\right) e_j(k), & \text{if } |e_j(k)| > \delta. \end{cases} \tag{14}$$

The deadzone width δ is chosen such that $\boldsymbol{\varphi}(v_j) \approx 0$ based on measured estimates of the iteration-varying disturbances, see Section 5.1. Deadzone gain γ acts as an additional learning gain that is only applied when $|e_j(k)| > \delta$. By choosing $\gamma \approx 1$ and linear gain $\alpha \approx 0$, the deadzone nonlinearity in (13) allows for fast convergence when the error is large, and limited disturbance amplification when the error is small.

The deadzone $\varphi(e_j(k))$ is a static nonlinearity that satisfies the incremental sector condition

$$0 \le \frac{\varphi(a) - \varphi(b)}{a - b} \le \gamma, \tag{15}$$

for scalars $a \neq b$. This property enables the analysis of the convergence of the nonlinear ILC algorithm, see Section 4.3. Although $\varphi(e_j)$ in the input update (13) is selected to be a deadzone nonlinearity in this paper, the results hold for any static nonlinearity that satisfies the incremental sector condition (15). This introduces an additional degree of freedom in the proposed nonlinear ILC scheme.

4.2. Convergence of nonlinear ILC

For linear ILC, monotonic convergence is typically analyzed for the case of $v_j = 0$. Due to linearity, (monotonic) convergence is also preserved for the case of non-zero v_j. For nonlinear ILC, as considered in this section, it is not sufficient to analyze convergence only for the case of $v_j = 0$. Even if a nonlinear system converges monotonically to a fixed point for $v_j = 0$, adding a (small) nonzero noise v_j can render the system unstable. Thus, for nonlinear ILC, (monotonic) convergence in the presence of non-zero v_j needs to be analyzed. This is not covered by the conventional definition of monotonic convergence from the case of linear ILC. Below, the notion of convergent (nonlinear) systems is introduced, that allows one to analyze (monotonic) convergence of nonlinear ILC for non-zero iteration-varying inputs v_j.

The convergence property of a (nonlinear) system implies that all solutions converge to a bounded steady-state solution. In particular, there exist conditions that ensure that a system is convergent for any bounded input, including non-constant and non-periodic inputs. Convergence of a discrete-time system is defined as follows (Pavlov & Van De Wouw, 2012, Definition 1).

Definition 3. A system

$$f_{j+1} = h(f_j, j), \qquad j \in \mathbb{Z} \tag{16}$$

is called exponentially convergent if

- there exists a unique solution $\bar{f}_j$ that is defined and bounded on $\mathbb{Z}$ (from $-\infty$ to $+\infty$); and
- $\bar{f}_j$ is globally exponentially stable, i.e., there exist $c > 0$ and $0 < \lambda < 1$ such that $|f_j - \bar{f}_j| \le c\lambda^{j-j_0}|f_{k_0} - \bar{f}_{j_0}|$ for all $j \ge j_0$.

The solution $\bar{f}_j$ is called a steady-state solution. Any solution of a convergent system converges to the steady-state solution, irrespective of the initial condition. Note that the time dependency on the right-hand side of system (16) is typically due to some input d_j, such that $f_{j+1} = h(f_j, d_j)$. The steady-state solution is determined only by the input d_j, and for a constant d_j, $\bar{f}_j$ is also constant, see, e.g., Pavlov and Van De Wouw (2012). For the convergence by Definition 3 of the (possibly nonlinear) system (16), the following holds (Pavlov & Van De Wouw, 2012, Theorem 1).

Lemma 4. *Consider system (16) with a right-hand side satisfying*

$$\|h(f^1, j) - h(f^2, j)\|_{\mathbf{P}} \le \lambda \|f^1 - f^2\|_{\mathbf{P}}, \qquad \forall f^1, f^2 \in \mathbb{R}^n, j \in \mathbb{Z} \tag{17}$$

$$\sup_{j \in \mathbb{Z}} \|h(0, j)\|_{\mathbf{P}} < +\infty, \tag{18}$$

for some matrix $\mathbf{P} = \mathbf{P}^{\top} \succ 0$ and $\lambda \in (0, 1)$. Then system (16) is exponentially convergent.

Stability of LTI systems, such as the feedforward iteration (4) of the linear ILC system, is a particular case of this convergence definition. This is shown in the following theorem, which also recovers the standard criterion for monotonic convergence of the sequence of iterates in the 2-norm, see Definition 5, as a special case.

Definition 5. A sequence of iterates $\{f_j\}$ is called monotonically convergent in the 2-norm to a unique steady-state solution $\bar{f}_j$ if

$$\|f_{j+1} - \bar{f}_{j+1}\|_2 \le \lambda \|f_j - \bar{f}_j\|_2, \ \forall j \in \mathbb{Z}, \tag{19}$$

for some $\lambda \in (0, 1)$.

Theorem 6. *The following statements hold for system (4) with any constant iteration-invariant input r and bounded sequence of iteration-varying disturbances v_j.*

(1) System (4) is exponentially convergent if

$$\|\mathbf{Q}(\mathbf{I} - \alpha\mathbf{L}\mathbf{J})\|_{\mathbf{P}} \le \lambda, \tag{20}$$

for some matrix $\mathbf{P} = \mathbf{P}^{\top} \succ 0$ and number $\lambda \in (0, 1)$.

(2) If (20) holds for $\mathbf{P} = \mathbf{I}$, i.e.,

$$\|\mathbf{Q}(\mathbf{I} - \alpha\mathbf{L}\mathbf{J})\|_{i2} = \bar{\sigma}(\mathbf{Q}(\mathbf{I} - \alpha\mathbf{L}\mathbf{J})) < 1, \tag{21}$$

then the solutions converge monotonically in the 2-norm to a unique steady-state solution that only depends on inputs r and v_j.

(3) If, in addition, $v_j = 0 \ \forall j$, then the sequence of iterates $\{f_j\}$ converges monotonically in the 2-norm to a fixed point f_∞:

$$\|f_{j+1} - f_\infty\|_2 \le \lambda \|f_j - f_\infty\|_2, \ \forall j \in \mathbb{Z}. \tag{22}$$

Proof. Regarding statement (1), consider two solutions f^1 and f^2 to (4), with the same iteration-varying input $\alpha\mathbf{Q}\mathbf{L}(r - v_j)$. Substituting in condition (17) gives

$$\begin{aligned} \left\|h(f^1, j) - h(f^2, j)\right\|_{\mathbf{P}} &= \big\|\mathbf{Q}(\mathbf{I} - \alpha\mathbf{L}\mathbf{J})f^1 + \alpha\mathbf{Q}\mathbf{L}(r - v_j) \\ &\quad - \mathbf{Q}(\mathbf{I} - \alpha\mathbf{L}\mathbf{J})f^2 - \alpha\mathbf{Q}\mathbf{L}(r - v_j)\big\|_{\mathbf{P}} \\ &= \|\mathbf{Q}(\mathbf{I} - \alpha\mathbf{L}\mathbf{J})(f^1 - f^2)\|_{\mathbf{P}} \\ &\le \|\mathbf{Q}(\mathbf{I} - \alpha\mathbf{L}\mathbf{J})\|_{\mathbf{P}} \|f^1 - f^2\|_{\mathbf{P}}. \end{aligned} \tag{23}$$

Thus, if (20) holds for some $\mathbf{P} = \mathbf{P}^\top \succ 0$ and $\lambda \in (0, 1)$ then also condition (17) is met, and the system (4) is exponentially convergent by Lemma 4. Statements (2) and (3) follow from substitution of $\mathbf{P} = \mathbf{I}$ and from the property that for convergent systems, a constant input leads to a constant steady-state solution, see, e.g., Pavlov and Van De Wouw (2012). □

For the LTI system (4), exponential convergence according to Theorem 6 is equivalent to exponential stability. The notion of monotonic convergence in the 2-norm of the sequence of iterates is common in ILC (Bristow et al., 2006) and ensures good learning transients (Longman, 2000). Typically, only the noise-free case is considered. Theorem 6 extends the notion of convergence in ILC to the noisy case, and the analysis of convergence in an arbitrary $\mathbf{P}$-norm connects standard ILC to the nonlinear ILC approach developed in this paper.

4.3. Criterion for convergence of nonlinear ILC

A criterion for the exponential convergence of nonlinear ILC is developed by interpreting the algorithm as a MIMO Lur'e system in the iteration domain. The criterion ensures that the system is convergent for any bounded input, including the combination of iteration-invariant and iteration-varying disturbances considered in this paper. The following auxiliary result enables convergence analysis of algorithm (13).

Lemma 7 (*Loop Transformation*)**.** *The nonlinear ILC update (13) with nonlinearity $\varphi(e_j(k))$ satisfying sector condition (15) is equivalent to*

$$f_{j+1} = \mathbf{Q}\left(f_j + \left(\alpha + \frac{\gamma}{2}\right)\mathbf{L}e_j + \mathbf{L}\tilde{\boldsymbol{\varphi}}(e_j)\right), \tag{24}$$

$$\tilde{\boldsymbol{\varphi}}(e_j) = \boldsymbol{\varphi}(e_j) - \frac{\gamma}{2}e_j \tag{25}$$

with $\tilde{\varphi}(e_j(k))$ satisfying the symmetric sector condition

$$-\frac{\gamma}{2} \le \frac{\tilde{\varphi}(a) - \tilde{\varphi}(b)}{a-b} \le \frac{\gamma}{2}. \tag{26}$$

Substituting (2) in (24) leads to

$$f_{j+1} = \left(\alpha + \frac{\gamma}{2}\right)\mathbf{QL}(r - v_j) + \mathbf{Q}\left(\mathbf{I} - \left(\alpha + \frac{\gamma}{2}\right)\mathbf{LJ}\right)f_j + \mathbf{QL}\tilde{\boldsymbol{\varphi}}(r - \mathbf{J}f_j - v_j). \tag{27}$$

System (27) can be described as an $N \times N$ Lur'e system, consisting of a linear system M with a static nonlinearity φ in feedback and external inputs d_1 and d_2, see Fig. 4. The system M with state f_j is given by

$$f_{j+1} = \mathbf{A}f_j + \mathbf{B}u_j + d_{1,j} \qquad y_j = \mathbf{C}f_j, \tag{28}$$

with system matrices $\mathbf{A} = \mathbf{Q}\left(\mathbf{I} - \left(\alpha + \frac{\gamma}{2}\right)\mathbf{LJ}\right)$, $\mathbf{B} = \mathbf{QL}$ and $\mathbf{C} = -\mathbf{J}$. The external inputs $d_{1,j}$ and $d_{2,j}$ are chosen as

$$d_{1,j} = \left(\alpha + \frac{\gamma}{2}\right)\mathbf{QL}(r - v_j), \tag{29}$$

$$d_{2,j} = r - v_j, \tag{30}$$

such that the input to $\tilde{\varphi}$ is $e_j = r - \mathbf{J}f_j - v_j$ and the output u_j of the feedback nonlinearity is given by

$$u_j = \tilde{\boldsymbol{\varphi}}(r - \mathbf{J}f_j - v_j). \tag{31}$$

The following theorem holds for the convergence, according to Definition 3, of the input iteration (27) of the nonlinear ILC system, which is equivalent to the Lur'e system (28).

Theorem 8. *The dynamic system defined by the input iteration (27) is exponentially convergent if the following conditions hold.*

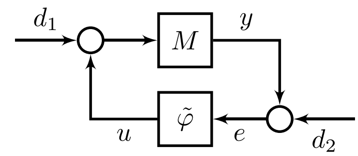

Fig. 4. Lur'e system with M a linear system and φ a static, memoryless nonlinearity.

A1. For matrix $\mathbf{A}$ *in (28) it holds that* $\rho(\mathbf{A}) < 1$, *i.e.,*

$$\max_i \left|\lambda_i\left(\mathbf{Q}\left(\mathbf{I} - \left(\alpha + \frac{\gamma}{2}\right)\mathbf{LJ}\right)\right)\right| < 1. \tag{32}$$

A2. Nonlinearity φ satisfies (15).
A3.

$$\sup_{\omega\in[0,2\pi)} \bar{\sigma} \left(\mathbf{J}\left(e^{i\omega}\mathbf{I} - \mathbf{Q}\left(\mathbf{I} - \left(\alpha + \frac{\gamma}{2}\right)\mathbf{LJ}\right)\right)^{-1}\mathbf{QL}\right) < \frac{2}{\gamma}. \tag{33}$$

The proof is given in the Appendix. Theorem 8 ensures that the nonlinear system is convergent for any input, including iteration-varying disturbances such as noise. For the noise-free case with $v_j = 0\ \forall j$, the external input of the Lur'e system is constant, and convergence of the system implies that the steady-state solution is also constant (Pavlov & Van De Wouw, 2012).

Theorem 8 provides a criterion for exponential convergence of the nonlinear ILC system. In typical ILC results, including Longman (2000), monotonic convergence in the 2-norm, see Definition 5, is desired to ensure good learning transients. For monotonic convergence in the 2-norm for nonlinear ILC, which requires that Lemma 4 holds for $\mathbf{P} = \mathbf{I}$, the following holds.

Corollary 9. *The dynamic system defined by the input iteration (27) is monotonically convergent in the 2-norm if*

B1.

$$\bar{\sigma}\left(\mathbf{Q}\left(\mathbf{I} - \left(\alpha + \frac{\gamma}{2}\right)\mathbf{LJ}\right)\right) < 1, \tag{34}$$

B2. φ satisfies (15), and
B3.

$$\begin{bmatrix} \mathbf{A}^\top\mathbf{A} + \mathbf{C}^\top\mathbf{C} - \mathbf{I} & \mathbf{A}^\top\mathbf{B} \\ \mathbf{B}^\top\mathbf{A} & \mathbf{B}^\top\mathbf{B} - \left(\frac{2}{\gamma}\right)^2\mathbf{I} \end{bmatrix} \prec 0, \tag{35}$$

with system matrices $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{C}$ *in (28).*

The proof is related to the proof of Theorem 8 and is given in the Appendix.

Remark 10. If the inverse $\mathbf{J}^{-1}$ exists, the system defined by the error iteration with a nonlinear ILC update can be constructed as follows.

$$e_{j+1} = \left(\mathbf{JQJ}^{-1} - \left(\alpha + \frac{\gamma}{2}\right)\mathbf{JQL}\right)e_j - \mathbf{JQL}\tilde{\boldsymbol{\varphi}}(e_j) + (\mathbf{I} - \mathbf{JQJ}^{-1})r + \mathbf{JQJ}^{-1}v_j - v_{j+1}. \tag{36}$$

This constitutes a Lur'e system similar to the input iteration, for which conditions for the convergence of error iteration (36) can be developed similar to Theorem 8. In addition, conditions for the convergence of the error iteration (5) for linear ILC can be developed similar to Theorem 6.

5. Design procedure

In this section, guidelines for the parameter selection are provided. The design procedure is summarized in Procedure 11 and is further described in the remainder of the section.

Procedure 11 Nonlinear ILC design

1: Determine deadzone width δ based on system measurements (Section 5.1).
2: Design learning and robustness matrices $\mathbf{Q}$ and $\mathbf{L}$ using standard lifted ILC methods (Section 5.2).
3: Select learning gains $\alpha \approx 0$ and $\gamma \approx 1$ while satisfying the convergence conditions of Theorem 8 (Section 5.3).

5.1. Deadzone selection

The aim is to attenuate repeating disturbances without amplifying the iteration-varying disturbances. Therefore, the deadzone width δ is determined based on the amplitude of the iteration-varying disturbances. The experimental data needed to determine δ is typically already available before ILC is implemented, because it follows from the standard operation of the system. Consider a series of n_e standard-operation experiments on the system (2) with a constant input $f_j = 0$ for $j = 1, 2, \ldots, n_e$. Each experiment gives

$$e_j = r - v_j, \tag{37}$$

with r the iteration-invariant part of the disturbances, and v_j a realization of the iteration-varying disturbances. Note that $E\{v_j\} = 0$ by definition, since all iteration-invariant disturbances are included in r. An estimate $\hat{r}$ of the iteration-invariant disturbance is obtained by computing the sample mean of the (37) over n_e experiments, see also Oomen (2020):

$$\hat{r} = \frac{1}{n_e}\sum_{j=0}^{n_e-1} e_j = r - \frac{1}{n_e}\sum_{j=0}^{n_e-1} v_j. \tag{38}$$

Then, for each experiment e_j an estimate of the iteration-varying disturbances is given by

$$\hat{v}_j = \hat{r} - e_j. \tag{39}$$

Based on the estimates $\hat{r}$ and $\hat{v}_j$, $j = 1, 2, \ldots, n_e$, a value of δ that filters out the desired percentage of iteration-varying disturbances is chosen. Note that for errors slightly larger than δ the gain is very small, because of the shape of the deadzone in Fig. 3. Therefore, disturbances that are slightly larger than δ are barely amplified, and δ does not have to fully encompass all iteration-varying disturbances. Fig. 5 shows an example of the estimates $\hat{r}$, $\hat{v}_j$ and a suitable choice of δ for the simulation example in Section 6 with $n_e = 20$.

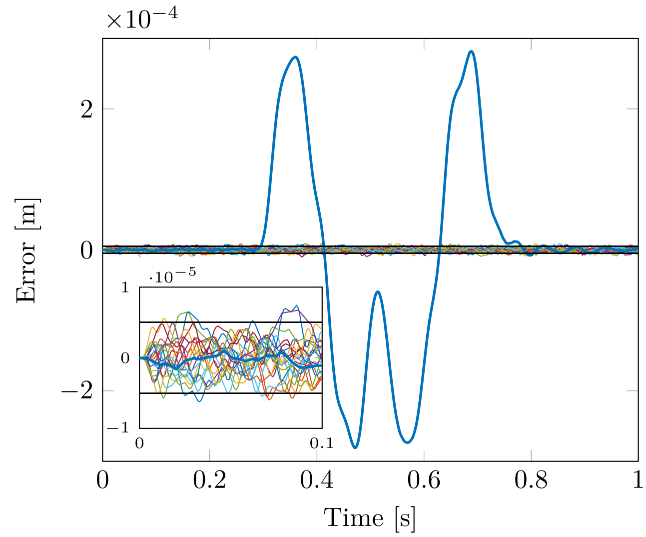

Fig. 5. Mean $\hat{r}$ of the error signal (—) and noise estimates $\hat{v}_j$, $j = 1, 2, \ldots, 20$ for 20 iterations. The interval $[-5 \times 10^{-6}, 5 \times 10^{-6}]$ (—) contains most of the noise.

5.2. Design of Q and L

The nonlinear ILC update law is designed based on norm-optimal ILC (Bristow et al., 2006; Gunnarsson & Norrlöf, 2001), yet instead of a weight on the change in input, a combination of linear and nonlinear learning gains is used. In standard norm-optimal ILC, the feedforward update is the model-based minimizer of a cost function of the form

$$\mathcal{J}(f_{j+1}) = \|e_{j+1}\|_{\mathbf{W}_e} + \|f_{j+1}\|_{\mathbf{W}_f} + \|f_{j+1} - f_j\|_{\mathbf{W}_{\Delta f}}. \tag{40}$$

This leads to the update (3) with matrices

$$\begin{aligned}\mathbf{Q} &= (\mathbf{J}^\top \mathbf{W}_e \mathbf{J} + \mathbf{W}_f + \mathbf{W}_{\Delta f})^{-1}(\mathbf{J}^\top \mathbf{W}_e \mathbf{J} + \mathbf{W}_{\Delta f}) \\ \mathbf{L} &= (\mathbf{J}^\top \mathbf{W}_e \mathbf{J} + \mathbf{W}_{\Delta f})^{-1}\mathbf{J}^\top \mathbf{W}_e,\end{aligned} \tag{41}$$

where the weight on the input $\mathbf{W}_f$ acts as a regularization term and the weight $\mathbf{W}_{\Delta f}$ on the change in input has a role similar to the learning gain α. For the nonlinear ILC algorithm, this last term is replaced by a combination of linear and nonlinear learning gains, leading to the ILC update

$$f_{j+1} = \mathbf{Q}(f_j + \alpha \mathbf{L} e_j + \mathbf{L}\varphi(e_j)) \tag{42}$$

with $\mathbf{Q}$ and $\mathbf{L}$ according to (41) with $\mathbf{W}_{\Delta f} = 0$. Note that if $\mathbf{J}$ is singular, $\mathbf{L}$ cannot be computed and the update should be implemented using $\mathbf{QL} = (\mathbf{J}^\top \mathbf{W}_e \mathbf{J} + \mathbf{W}_f)^{-1}\mathbf{J}^\top \mathbf{W}_e$. Taking $\mathbf{W}_e \succ 0$ and $\mathbf{W}_f \succeq 0$, ensures convergence in the linear case with non-singular $\mathbf{J}$ and a perfect model (Gunnarsson & Norrlöf, 2001). In practice $\mathbf{W}_f \succ 0$ is often required, both because $\mathbf{J}$ is often singular and to provide robustness against model uncertainty (van de Wijdeven, Donkers, & Bosgra, 2009). For $\mathbf{W}_e = I$, $\mathbf{W}_f$ should be chosen as small as possible but such that $\bar{\sigma}(\mathbf{Q}(\mathbf{I} - \mathbf{LJ})) < 1$, i.e., such that Condition (21) for monotonic convergence in the 2-norm of Theorem 6 is met for $\alpha = 1$, $\gamma = 0$.

5.3. Gain selection

The main idea of using a nonlinear learning filter in ILC is to ignore iteration-varying disturbances v_j completely, and achieve fast learning only for the iteration-invariant disturbance r. This suggests choosing the linear learning gain α in (13) approximately zero and the nonlinear gain $\gamma = 1$. However, iteration-varying disturbances may occur on top of iteration-invariant disturbances. Initially, these varying disturbances are amplified by γ, but as the iteration-invariant disturbances are attenuated the total error approaches the bounds of the deadzone and the influence of γ reduces. At this point the convergence speed and converged error depend on α, and for small α the initial amplification of the iteration-varying disturbances is reduced through averaging. It follows that α should be small but nonzero, while γ should be close to 1. The influence of γ and α on the convergence according to Theorem 8 should also be taken into account.

6. Example

In this section, linear and nonlinear ILC are compared in simulation. The system P is a 20th-order model of the carriage of an

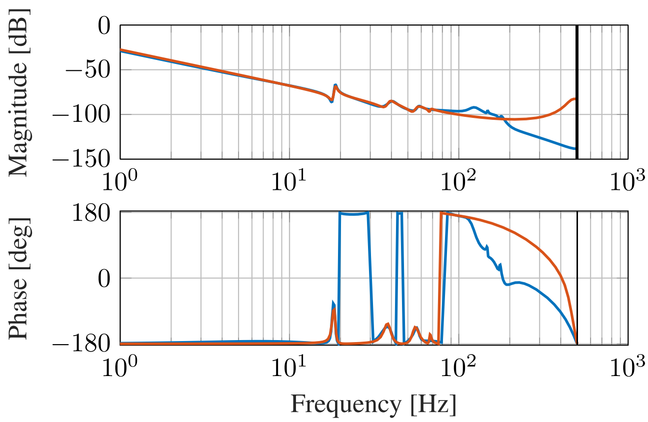

Fig. 6. Bode diagram of the system P (—) and a low-order approximation (—) that is used to design learning filter $\mathbf{L}$.

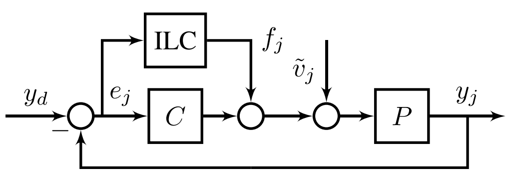

Fig. 7. Parallel ILC configuration.

industrial flatbed printer with a *PD*-type controller C in closed-loop. For the design of $\mathbf{L}$ and $\mathbf{Q}$, the system is approximated by a 12th-order model, resulting in a model mismatch at high frequencies, see Fig. 6. Parallel ILC is applied to this system, where f_j is a feedforward signal injected between P and C and $J = \frac{P}{1+PC}$, see Fig. 7. The iteration-invariant disturbance is given by $r = Sy_d$ with $S = \frac{1}{1+PC}$, and reference y_d a forward and backward motion with length $N = 1000$. The iteration-varying disturbance is given by $v_j = SP\tilde{v}_j$, where $\tilde{v}_j$ is Gaussian i.i.d. noise with a variance of 0.005 that is injected between the controller and the plant.

The parameters are chosen according to Procedure 11. Because of the model mismatch at high frequencies, $\mathbf{W}_f \succ 0$ is required and the weights are chosen as $\mathbf{W}_e = \mathbf{I}$ and $\mathbf{W}_f = 10^{-10}\mathbf{I}$. The scaling of $\mathbf{W}_e$ and $\mathbf{W}_f$ depends on the magnitudes of e and f. Based on Fig. 5, a deadzone width of $\delta = 5 \times 10^{-6}$ is chosen. To meet the convergence conditions in Theorem 8, the gains are chosen as $\gamma = 0.9$ and $\alpha = 0.1$.

The results in Fig. 8, which are averaged over 20 realizations, show that nonlinear ILC achieves both fast convergence and small converged errors. The convergence speed is comparable to the linear case with $\alpha = 1$, while the converged error 2-norm matches that for the linear case with $\alpha = 0.2$. This illustrates that nonlinear ILC can remove the trade-off between convergence speed and converged error.

7. Conclusions

A nonlinear ILC approach is introduced that achieves both fast convergence and small errors in the presence of iteration-varying disturbances. Through a deadzone nonlinearity, the traditional trade-off between convergence speed and amplification of iteration-varying disturbances is removed by applying various learning gains to different elements of the error signal depending on their magnitude. A condition for monotonic convergence is developed by interpreting the system as a Lur'e system in the iteration domain. In addition, a design procedure is provided based on disturbance measurements and system knowledge. The approach is validated in simulations in which fast convergence to small errors is demonstrated. Future work concerns experimental validation, as well as the analysis of the effect of the design parameters in nonlinear ILC on the expected value of the converged error, similar to the result in Theorem 2 for the linear case.

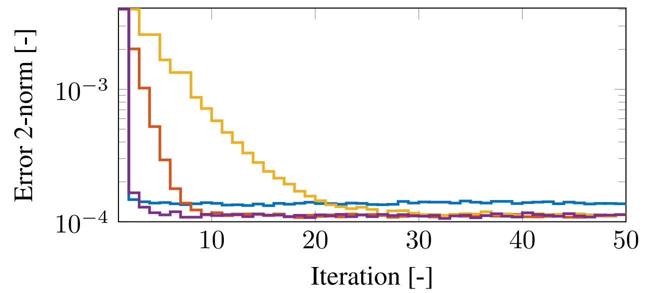

Fig. 8. Error 2-norm averaged over 20 realizations for lifted linear ILC with $\mathbf{W}_e = \mathbf{I}$, $\mathbf{W}_f = 10^{-10}\mathbf{I}$ and $\alpha = 1$ (—), 0.5 (—) and 0.2 (—) and nonlinear lifted ILC with $\alpha = 0.1$, $\gamma = 0.9$ and $\delta = 5 \times 10^{-6}$ (—). For linear ILC, small learning gains lead to smaller converged errors and reduced convergence speed. Nonlinear ILC removes this trade-off between convergence speed and converged error.

Appendix. Proofs of Theorems 2 and 8 and Corollary 9

In this appendix, the proofs of Theorems 2 and 8 and Corollary 9 are given.

Proof of Theorem 2. Consider the error iteration (5). By linearity of the expectation operator,

$$\begin{aligned} E\{e_{j+1}\} =& \mathbf{JQ}(\mathbf{J}^{-1} - \alpha\mathbf{L})E\{e_j\} \\ &+ (\mathbf{I} - \mathbf{JQJ}^{-1})r + \mathbf{JQJ}^{-1}E\{v_j\} - E\{v_{j+1}\} \\ =& \mathbf{JQ}(\mathbf{J}^{-1} - \alpha\mathbf{L})E\{e_j\} + (\mathbf{I} - \mathbf{JQJ}^{-1})r. \end{aligned} \tag{A.1}$$

Since $\|\mathbf{JQ}(\mathbf{J}^{-1} - \alpha\mathbf{L})\|_{\mathbf{P}} < \lambda$ for some $\mathbf{P} = \mathbf{P}^\top \succ 0$, $\lambda \in (0, 1)$, this system is exponentially stable and the constant input $(\mathbf{I} - \mathbf{JQJ}^{-1})r$ leads to a unique constant solution $E\{e_\infty\}$, given by

$$\begin{aligned} E\{e_\infty\} &= \mathbf{JQ}(\mathbf{J}^{-1} - \alpha\mathbf{L})E\{e_\infty\} + (\mathbf{I} - \mathbf{JQJ}^{-1})r \\ &= (\mathbf{I} - \mathbf{JQ}(\mathbf{J}^{-1} - \alpha\mathbf{L}))^{-1}(\mathbf{I} - \mathbf{JQJ}^{-1})r \end{aligned} \tag{A.2}$$

Next, consider the deviation $\tilde{e}_{j+1} = E\{e_{j+1}\} - e_{j+1}$, which is given by

$$\tilde{e}_{j+1} = \mathbf{JQ}(\mathbf{J}^{-1} - \alpha\mathbf{L})\tilde{e}_j - \mathbf{JQJ}^{-1}v_j + v_{j+1}. \tag{A.3}$$

Then,

$$\begin{aligned} \tilde{e}_{j+1}^\top\tilde{e}_{j+1} =& (\mathbf{JQJ}^{-1}v_j)^\top\mathbf{JQJ}^{-1}v_j + v_{j+1}^\top v_{j+1} \\ &+ (\mathbf{JQ}(\mathbf{J}^{-1} - \alpha\mathbf{L})\tilde{e}_j)^\top(\mathbf{JQ}(\mathbf{J}^{-1} - \alpha\mathbf{L})\tilde{e}_j) \\ &+ 2[-(\mathbf{JQJ}^{-1}v_j)^\top v_{j+1} - (\mathbf{JQJ}^{-1}v_j)^\top\mathbf{JQ}(\mathbf{J}^{-1} - \alpha\mathbf{L})\tilde{e}_j \\ &+ v_{j+1}^\top\mathbf{JQ}(\mathbf{J}^{-1} - \alpha\mathbf{L})\tilde{e}_j]. \end{aligned} \tag{A.4}$$

The variance $E\{\tilde{e}_{j+1}^\top\tilde{e}_{j+1}\}$ of e_{j+1} is then given by

$$\begin{aligned} E\{\tilde{e}_{j+1}^\top\tilde{e}_{j+1}\} =& E\{(\mathbf{JQJ}^{-1}v_j)^\top\mathbf{JQJ}^{-1}v_j\} + E\{v_{j+1}^\top v_{j+1}\} \\ &+ E\{(\mathbf{JQ}(\mathbf{J}^{-1} - \alpha\mathbf{L})\tilde{e}_j)^\top(\mathbf{JQ}(\mathbf{J}^{-1} - \alpha\mathbf{L})\tilde{e}_j)\} \\ &- 2E\{(\mathbf{JQJ}^{-1}v_j)^\top\mathbf{JQ}(\mathbf{J}^{-1} - \alpha\mathbf{L})\tilde{e}_j\}, \end{aligned} \tag{A.5}$$

since $E\{v_jv_{j+1}\} = 0$ and $E\{v_{j+1}\tilde{e}_j\} = 0$ due to Assumption 1. For the last cross-term, $\tilde{e}_j$ follows from evaluating (A.3) at iteration j

and is premultiplied by $(\mathbf{JQJ}^{-1}v_j)^\mathsf{T}\mathbf{JQ}(\mathbf{J}^{-1}-\alpha\mathbf{L})$, hence

$$E\{(\mathbf{JQJ}^{-1}v_j)^\mathsf{T}\mathbf{JQ}(\mathbf{J}^{-1}-\alpha\mathbf{L})\tilde{e}_j\} \\ = E\{(\mathbf{JQJ}^{-1}v_j)^\mathsf{T}\mathbf{JQ}(\mathbf{J}^{-1}-\alpha\mathbf{L})v_j\}. \tag{A.6}$$

Therefore,

$$E\{\tilde{e}_{j+1}^\mathsf{T}\tilde{e}_{j+1}\} = E\{(\mathbf{JQJ}^{-1}v_j)^\mathsf{T}\mathbf{JQJ}^{-1}v_j\} + E\{v_{j+1}^\mathsf{T}v_{j+1}\} \\ + E\{(\mathbf{JQ}(\mathbf{J}^{-1}-\alpha\mathbf{L})\tilde{e}_j)^\mathsf{T}(\mathbf{JQ}(\mathbf{J}^{-1}-\alpha\mathbf{L})\tilde{e}_j)\} \\ - 2E\{(\mathbf{JQJ}^{-1}v_j)^\mathsf{T}\mathbf{JQ}(\mathbf{J}^{-1}-\alpha\mathbf{L})v_j\}, \tag{A.7}$$

which concludes the proof. □

The proof of Theorem 8 employs the discrete-time Kalman–Yakubovich–Popov lemma (see, e.g., Rantzer (1996, Theorem 2)) in the following form.

Lemma 12. *Given* $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$, *and* $\mathbf{M} = \mathbf{M}^\mathsf{T} \in \mathbb{R}^{(n+m)\times(n+m)}$, *with* $\rho(\mathbf{A}) < 1$, *the following two statements are equivalent:*

(1)

$$\begin{bmatrix}(e^{j\omega}\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\\ \mathbf{I}\end{bmatrix}^* \mathbf{M} \begin{bmatrix}(e^{j\omega}\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\\ \mathbf{I}\end{bmatrix} \prec 0 \forall\omega\in\mathbb{R} \tag{A.8}$$

(2) There exists a matrix $\mathbf{P} = \mathbf{P}^\mathsf{T} \in \mathbb{R}^{n\times n}$ *such that*

$$\mathbf{M} + \begin{bmatrix}\mathbf{A}^\mathsf{T}\mathbf{PA}-\mathbf{P} & \mathbf{A}^\mathsf{T}\mathbf{PB}\\ \mathbf{B}^\mathsf{T}\mathbf{PA} & \mathbf{B}^\mathsf{T}\mathbf{PB}\end{bmatrix} \prec 0, \tag{A.9}$$

and if the upper left corner of $\mathbf{M}$ *is positive semidefinite, it follows from Schur stability of* $\mathbf{A}$ *that* $\mathbf{P} \succ 0$.

Proof of Theorem 8. First, condition A1 ensures that the linear system is exponentially stable, ensuring exponential convergence for any bounded input for which $\tilde{\boldsymbol{\varphi}}(e_j) = 0\ \forall j$. Secondly, conditions A1–A3 ensure that the nonlinear Lur'e system (28) is exponentially convergent for any bounded input. The condition of Lemma 4 is applied to (28). Take

$$\begin{aligned} h(f^1,j)-h(f^2,j) =& \mathbf{A}f^1 + \mathbf{B}\tilde{\boldsymbol{\varphi}}(Cf^1+d_{2,j}) + d_{1,j} - \\ & \mathbf{A}f^2 - \mathbf{B}\tilde{\boldsymbol{\varphi}}(Cf^2+d_{2,j}) - d_{1,j} \\ =& \mathbf{A}\Delta f(k) + \mathbf{B}\Delta u(k), \end{aligned} \tag{A.10}$$

with $\Delta f = f^1 - f^2$ and $\Delta u = \Delta\tilde{\boldsymbol{\varphi}} = \tilde{\boldsymbol{\varphi}}(Cf^1+d_{2,j}) - \tilde{\boldsymbol{\varphi}}(Cf^2+d_{2,j})$. Using (A.10), condition (17) is rewritten to

$$\begin{bmatrix}\Delta f\\ \Delta u\end{bmatrix}^\mathsf{T} \begin{bmatrix}\mathbf{A}^\mathsf{T}\mathbf{PA} & \mathbf{A}^\mathsf{T}\mathbf{PB}\\ \mathbf{B}^\mathsf{T}\mathbf{PA} & \mathbf{B}^\mathsf{T}\mathbf{PB}\end{bmatrix} \begin{bmatrix}\Delta f\\ \Delta u\end{bmatrix} \\ - \lambda^2\Delta f^\mathsf{T}\mathbf{P}\Delta f < 0,\ \forall\Delta f, \Delta u\in\mathbb{R}^n, j\in\mathbb{Z}. \tag{A.11}$$

Next, the symmetric sector condition for $\tilde{\boldsymbol{\varphi}}$ is written as a quadratic matrix inequality. Denote $z^i = \mathbf{C}f^i + d_{2,j}$ and $\Delta z = z^1 - z^2$. By condition A2, $\tilde{\boldsymbol{\varphi}}$ satisfies the symmetric sector condition with $\frac{\gamma}{2}$, and therefore

$$\begin{bmatrix}\Delta z\\ \Delta\tilde{\boldsymbol{\varphi}}\end{bmatrix}^\mathsf{T} \begin{bmatrix}(\frac{\gamma}{2})^2\mathbf{I} & 0\\ 0 & -\mathbf{I}\end{bmatrix} \begin{bmatrix}\Delta z\\ \Delta\tilde{\boldsymbol{\varphi}}\end{bmatrix} \geq 0. \tag{A.12}$$

Substituting $\Delta z = \mathbf{C}\Delta f$, $\Delta\tilde{\boldsymbol{\varphi}} = \Delta u$, and dividing by $(\frac{\gamma}{2})^2$ gives

$$\begin{bmatrix}\Delta f\\ \Delta u\end{bmatrix}^\mathsf{T} \begin{bmatrix}\mathbf{C}^\mathsf{T}\mathbf{C} & 0\\ 0 & -\left(\frac{2}{\gamma}\right)^2\mathbf{I}\end{bmatrix} \begin{bmatrix}\Delta f\\ \Delta u\end{bmatrix} \geq 0. \tag{A.13}$$

The system is exponentially convergent if there exists a $\mathbf{P} = \mathbf{P}^\mathsf{T} \succ 0$ and $\lambda \in (0,1)$ such that (A.11) holds for all $\Delta f, \Delta u$ that satisfy (A.13). By the S-procedure (Pólik & Terlaky, 2007), this holds if there exist a $\mathbf{P} = \mathbf{P}^\mathsf{T} \succ 0$ and $\lambda \in (0,1)$ such that

$$\begin{bmatrix}\Delta f\\ \Delta u\end{bmatrix}^\mathsf{T} \begin{bmatrix}\mathbf{A}^\mathsf{T}\mathbf{PA}+\mathbf{C}^\mathsf{T}\mathbf{C}-\lambda^2\mathbf{P} & \mathbf{A}^\mathsf{T}\mathbf{PB}\\ \mathbf{B}^\mathsf{T}\mathbf{PA} & \mathbf{B}^\mathsf{T}\mathbf{PB}-\left(\frac{2}{\gamma}\right)^2\mathbf{I}\end{bmatrix} \begin{bmatrix}\Delta f\\ \Delta u\end{bmatrix} \\ \leq 0,\ \forall\Delta f, \Delta u\in\mathbb{R}^n, j\in\mathbb{Z}. \tag{A.14}$$

Condition A3 implies the existence of a $\mathbf{P} = \mathbf{P}^\mathsf{T} \succ 0$ and $\lambda \in (0,1)$ such that (A.14) holds, as is shown next. Substitute $\mathbf{A} = \mathbf{Q}\left(\mathbf{I} - \left(\alpha + \frac{\gamma}{2}\right)\mathbf{LJ}\right)$, $\mathbf{B} = \mathbf{QL}$ and $\mathbf{C} = -\mathbf{J}$ according to (28) in Condition A3 and rewrite to obtain

$$\begin{bmatrix}(e^{i\omega}\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\\ \mathbf{I}\end{bmatrix}^* \begin{bmatrix}\mathbf{C}^\mathsf{T}\mathbf{C} & 0\\ 0 & -\left(\frac{2}{\gamma}\right)^2\mathbf{I}\end{bmatrix} \begin{bmatrix}(e^{i\omega}\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\\ \mathbf{I}\end{bmatrix} < 0 \\ \forall\omega\in\mathbb{R}, \tag{A.15}$$

which is equivalent to (A.8) in Lemma 12 with

$$\mathbf{M} = \begin{bmatrix}\mathbf{C}^\mathsf{T}\mathbf{C} & 0\\ 0 & -\left(\frac{2}{\gamma}\right)^2\mathbf{I}\end{bmatrix}. \tag{A.16}$$

Therefore, by Lemma 12, Condition A3 implies that there exists a matrix $\mathbf{P} = \mathbf{P}^\mathsf{T} \succ 0$ such that

$$\begin{bmatrix}\mathbf{A}^\mathsf{T}\mathbf{PA}+\mathbf{C}^\mathsf{T}\mathbf{C}-\mathbf{P} & \mathbf{A}^\mathsf{T}\mathbf{PB}\\ \mathbf{B}^\mathsf{T}\mathbf{PA} & \mathbf{B}^\mathsf{T}\mathbf{PB}-\left(\frac{2}{\gamma}\right)^2\mathbf{I}\end{bmatrix} < 0, \tag{A.17}$$

and therefore also

$$\begin{bmatrix}\Delta f\\ \Delta u\end{bmatrix}^\mathsf{T} \begin{bmatrix}\mathbf{A}^\mathsf{T}\mathbf{PA}+\mathbf{C}^\mathsf{T}\mathbf{C}-\mathbf{P} & \mathbf{A}^\mathsf{T}\mathbf{PB}\\ \mathbf{B}^\mathsf{T}\mathbf{PA} & \mathbf{B}^\mathsf{T}\mathbf{PB}-\left(\frac{2}{\gamma}\right)^2\mathbf{I}\end{bmatrix} \begin{bmatrix}\Delta f\\ \Delta u\end{bmatrix} < 0 \\ \forall\Delta f, \Delta u\in\mathbb{R}^n. \tag{A.18}$$

Since (A.18) holds, there always exists a value $\varepsilon \in (0,1)$ small enough such that

$$\begin{bmatrix}\Delta f\\ \Delta u\end{bmatrix}^\mathsf{T} \begin{bmatrix}\mathbf{A}^\mathsf{T}\mathbf{PA}+\mathbf{C}^\mathsf{T}\mathbf{C}-\mathbf{P} & \mathbf{A}^\mathsf{T}\mathbf{PB}\\ \mathbf{B}^\mathsf{T}\mathbf{PA} & \mathbf{B}^\mathsf{T}\mathbf{PB}-\left(\frac{2}{\gamma}\right)^2\mathbf{I}\end{bmatrix} \begin{bmatrix}\Delta f\\ \Delta u\end{bmatrix} \\ + \Delta f^\mathsf{T}\varepsilon\mathbf{P}\Delta f \leq 0,\ \forall\Delta f, \Delta u\in\mathbb{R}^n. \tag{A.19}$$

It follows that Condition A3 ensures the existence of $\lambda = \sqrt{1-\varepsilon} \in (0,1)$ and $\mathbf{P} = \mathbf{P}^\mathsf{T} \succ 0$ for which (A.14) holds. Therefore, if conditions A1–A3 are satisfied, the system (28) is exponentially convergent for any bounded input. □

Next, the proof of Corollary 9 is given.

Proof of Corollary 9. Condition B1 ensures that the error iteration $\{e_j\}$ of the linear system is monotonically convergent in the 2-norm according to Theorem 6 when $\tilde{\boldsymbol{\varphi}}(e) = 0$. Next, conditions B2 and B3 ensure that Lemma 4 holds for $\mathbf{P} = \mathbf{I}$, which follows from the proof of Theorem 8, cf. Eq. (A.18) with $\mathbf{P} = \mathbf{I}$. □

References

Aarnoudse, Leontine, Pavlov, Alexey, & Oomen, Tom (2023). Nonlinear iterative learning control: a frequency-domain approach for fast convergence and high accuracy. In *2023 IFAC world congr.* (pp. 1889–1894). Yokohama, Japan.

Ahn, Hyun-Sik, Choi, Chong-Ho, & Kim, Kwang-Bae (1993). Iterative learning control for a class of nonlinear systems. *Automatica*, *29*(6), 1575–1578.

Altın, Berk, & Barton, Kira (2017). Exponential stability of nonlinear differential repetitive processes with applications to iterative learning control. *Automatica*, *81*, 369–376.

Bristow, D. A., Tharayil, M., & Alleyne, A. G. (2006). A survey of iterative learning control. *IEEE Control Systems*, *26*(3), 96–114.

Butcher, M., Karimi, A., & Longchamp, R. (2008). A statistical analysis of certain iterative learning control algorithms. *International Journal of Control*, *81*(1), 156–166.

Gunnarsson, Svante, & Norrlöf, Mikael (2001). On the design of ILC algorithms using optimization. *Automatica*, *37*(12), 2011–2016.

Heertjes, Marcel, Rampadarath, Randjanie, & Waiboer, Rob (2009). Nonlinear Q-filter in the learning of nano-positioning motion systems. In *2009 eur. control conf.* (pp. 1523–1528). Budapest, Hungary.

Heertjes, Marcel, & Steinbuch, Maarten (2004). Stability and performance of a variable gain controller with application to a dvd storage drive. *Automatica*, *40*(4), 591–602.

Heertjes, Marcel, & Tso, Tim (2007). Robustness, convergence, and Lyapunov stability of a nonlinear iterative learning control applied at a wafer scanner. In *Proc. 2007 am. control conf.* (pp. 5490–5495). New York City, USA.

Longman, Richard W. (2000). Iterative learning control and repetitive control for engineering practice. *International Journal of Control*, *73*(10), 930–954.

Longman, Richard W. (2010). On the theory and design of linear repetitive control systems. *European Journal of Control*, *16*(5), 447–496.

Oomen, Tom (2020). Learning for advanced motion control. In *IEEE int. work. adv. motion control.* Agder, Norway.

Oomen, Tom, & Rojas, Cristian R. (2017). Sparse iterative learning control with application to a wafer stage: Achieving performance, resource efficiency, and task flexibility. *Mechatronics*, *47*, 134–147.

Pavlov, A., Hunnekens, B. G. B., Wouw, N. V. D., & Nijmeijer, H. (2013). Steady-state performance optimization for nonlinear control systems of Lur'e type. *Automatica*, *49*(7), 2087–2097.

Pavlov, Alexey, & Van De Wouw, Nathan (2012). Steady-state analysis and regulation of discrete-time nonlinear systems. *IEEE Transactions on Automatic Control*, *57*(7), 1793–1798.

Pólik, Imre, & Terlaky, Tamás (2007). A survey of the S-lemma. *SIAM Review*, *49*(3), 371–418.

Rantzer, Anders (1996). On the Kalman—Yakubovich—Popov lemma. *Systems & Control Letters*, *28*(1), 7–10.

van de Wijdeven, Jeroen, Donkers, Tijs, & Bosgra, Okko (2009). Iterative Learning Control for uncertain systems: Robust monotonic convergence analysis. *Automatica*, *45*(10), 2383–2391.

Volckaert, Marnix, Diehl, Moritz, & Swevers, Jan (2013). Generalization of norm optimal ILC for nonlinear systems with constraints. *Mechanical Systems and Signal Processing*, *39*(1–2), 280–296.

Leontine Aarnoudse received the B.Sc degree (2017), M.Sc degree (cum laude) (2019) and Ph.D. degree (2024) in Mechanical Engineering from the Eindhoven University of Technology, Eindhoven, The Netherlands. She is currently affiliated with the Norwegian University of Science and Technology as a postdoctoral researcher. Her research interests are in the field of control for precision mechatronics, and are mostly centered around the development of learning theory for these systems.

Alexey Pavlov obtained his M.Sc.-degree (cum laude) in Applied Mathematics from St. Petersburg State University, Russia, and Ph.D.-degree in Mechanical Engineering from Eindhoven University of Technology, The Netherlands in 1998 and 2004, respectively. Currently he holds a full professor position at the Department of Geoscience and Petroleum of the Norwegian University of Science and Technology (NTNU). He held various academic and industrial positions at the Institute for Problems of Mechanical Engineering, Russia (1999), Ford Research Laboratory, USA (2000), Eindhoven University of Technology (2005), NTNU (2005–2009), Equinor Research Centre, Norway (2009–2016). He published the book 'Uniform Output Regulation of Nonlinear Systems: A convergent Dynamics Approach' with N. van de Wouw and H. Nijmeijer (Birkhauser, 2005). He received the IEEE Control Systems Technology Award (2015) "For the development and application of variable-gain control techniques for high-performance motion systems" and IFAC World Congress Best Application Paper Award (2011) for the work "Drilling seeking automatic control solutions". His inventions within automatic optimization are implemented in several oil production fields. His current research interests include control of nonlinear systems, control and optimization of highly uncertain systems, automation and digitalization in the energy industry and various industrial applications of automatic control and optimization.

Tom Oomen is full professor with the Department of Mechanical Engineering at the Eindhoven University of Technology. He is also a part-time full professor with the Delft University of Technology. He received the M.Sc. degree (cum laude) and Ph.D. degree from the Eindhoven University of Technology, Eindhoven, The Netherlands. He held visiting positions at KTH, Stockholm, Sweden, and at The University of Newcastle, Australia. He is a recipient of the 7th Grand Nagamori Award, the Corus Young Talent Graduation Award, the IFAC 2019 TC 4.2 Mechatronics Young Research Award, the 2015 IEEE Transactions on Control Systems Technology Outstanding Paper Award, the 2017 IFAC Mechatronics Best Paper Award, the 2019 IEEJ Journal of Industry Applications Best Paper Award, and recipient of a Veni and Vidi personal grant. He is currently a Senior Editor of IEEE Control Systems Letters (L-CSS) and Associate Editor of IFAC Mechatronics, and he has served on the editorial boards of the IEEE Control Systems Letters (L-CSS) and IEEE Transactions on Control Systems Technology. He has also been vice-chair for IFAC TC 4.2 and a member of the Eindhoven Young Academy of Engineering. His research interests are in the field of data-driven modeling, learning, and control, with applications in precision mechatronics.