Abstract—This paper studies optimal design of Iterative Learning (ILC) and Repetitive Control (RC) for linear, cycle invariant systems. The optimization is based on minimizing the 2-norm of time signals on a repeating time interval. In contrast to weighting the error and actuating variable, as generally proposed, a different interpretation of the weighting is suggested. This is based on the similarity of 2-norm-optimal designs to Tikhonov regularization, which also makes optimal designs well suited for ill-conditioned systems. The choice of weightings is discussed in comparison to well known ILC designs. Finally, a novel implementation of ILC/RC saving memory and calculation time is presented.

I. INTRODUCTION

Iterative Learning Control (ILC) and Repetitive Control (RC) both aim to perfectly track a desired finite time trajectory, which in the following is referred to as a cycle. The basic idea is to enhance control performance of each cycle by utilizing information of previous ones. Whereas in ILC the cycles are modeled to be starting and ending in the same states, in RC the cycles follow each other consecutively. Thus, the initial states of a cycle are determined by the end states of the previous cycle. Basic concepts and stability proofs bridge ILC and RC [14]. However, a differentiation between ILC and RC is of interest in certain applications. Major differences have been explored in [13].

ILC/RC has a two dimensional structure. In its first dimension ILC/RC is a time domain feed forward control. In its second dimension it is a cycle domain feedback control. Some publications consider both dimensions for analysis and controller design [20], [22], [8]. Often only the second dimension, the cycle domain feedback, is considered to be essential part of ILC/RC [9].

Generally ILC/RC concepts have a fixed cycle domain feedback structure of which the P- and D-type are the most famous ones [2], [12]. These designs are limited in transient behavior and stable convergence. Vector notation of the ILC/RC problems provides a general framework, which enables to formulate variable types of learning laws [19], [4], [15]. Therefore, vector notation of ILC/RC has gained popularity in recent years. A drawback of vector notation is its calculation and memory intense implementation.

Optimal designs of ILC/RC problems are based on minimizing a cost function, which contains the weighted control error and actuating variable [21], [5]. Vector notation enables to algebraically describe the entire error and actuating signal over the cycle. It is therefore well suited for optimal designs as investigated in the past [1], [11], [6], [17], [4]. Optimal design assures monotonic and stable convergence. This holds true even for ill-conditioned and non-minimum phase systems.

Purpose of this paper is to present the idea, advantages and limits of optimal ILC/RC designs consolidating previous works. Most notably this paper suggests a different interpretation of the weighting parameters, which allows a decided choice without need of heuristic tuning. Concluding, a novel implementation of optimal ILC/RC is addressed. Optimal designs allow a reduced implementation of ILC/RC. This saves calculation time and memory without loss of any degree of freedom in controller design.

II. PROBLEM FORMULATION

Consider the linear, discrete SISO system with input \( u(i) \), output \( y(i) \) yielding the following state space model

\[
\begin{align*}
\mathbf{x}(i + 1) &= \mathbf{A}\mathbf{x}(i) + \mathbf{b}u(i) \\
y(i) &= \mathbf{c}\mathbf{x}(i)
\end{align*}
\]

(1)

with initial states \( \mathbf{x}(0) = 0 \) and sampling time \( T_s \). The state space model does not have feed through like most mechanical systems. Lifting (1) on the repeating cycle of length \( T_r \) leads to the vector notation

\[
y = \mathbf{G}u
\]

(2)

with the sampled data vectors

\[
\begin{align*}
\mathbf{u} &= [u(0), u(1), \ldots, u(i), \ldots, u(N-1)]^T, \\
\mathbf{y} &= [y(0), y(1), \ldots, y(i), \ldots, y(N-1)]^T
\end{align*}
\]

(3)

of length \( N = \frac{T_r}{T_s} \) and Toeplitz matrix

\[
\mathbf{G}_{ILC} = \begin{bmatrix}
0 & 0 & \cdots & 0 \\
g_1 & 0 & 0 & \vdots \\
\vdots & \ddots & \ddots & 0 \\
g_{N-1} & \cdots & g_1 & 0
\end{bmatrix}
\]

(4)

holding the Markov parameters \( g_i = \beta^{T_i-1} \Delta^{-1} \beta \) representing the discrete impulse response of the system. For RC and frequency based ILC designs a steady state system description
yields the cyclic matrix
\[
G_{RC} = \begin{bmatrix}
0 & g_{N-1} & \cdots & g_1 \\
g_1 & 0 & \ddots & \\
\vdots & \ddots & \ddots & \ddots \\
g_{N-1} & \cdots & g_1 & 0
\end{bmatrix}
\] (5)
as presented in [13].

The ILC/RC problem is stated in form of a MIMO \( R^N \) control problem in discrete cycle domain \( k \)
\[
\begin{align*}
\mathbf{u}_{k+1} &= S \mathbf{u}_k + \Gamma (\mathbf{w} - \mathbf{y}_k) \\
\mathbf{y}_k &= G \mathbf{u}_k
\end{align*}
\] (6)
with reference input vector \( \mathbf{w} \) as in (3) holding the sampled reference cycle. The matrix \( \Gamma \), which is to be determined in the controller design, describes the learning update. Matrix \( S \) has been introduced as a "memory", determining how much information of \( \mathbf{u}_k \) will be recalled in cycle \( k + 1 \). Zero error convergence can only be achieved if the memory matrix equals the identity matrix \( S = I \) [19].

Aim of ILC/RC is to find an optimal actuating vector \( \mathbf{u}_k \) fulfilling
\[
\lim_{k \to \infty} \mathbf{y}_k = G \mathbf{u}_\infty \approx \mathbf{w} \quad \text{or} \quad \lim_{k \to \infty} (\mathbf{w} - \mathbf{y}_k) = e \approx 0
\]
in a best possible approximation, by iteratively updating the previous actuating variable \( \mathbf{u}_k \) with the increment \( \Delta \mathbf{u}_k = \mathbf{r}_k \) as described in (6).

The ideal learning matrix \( \Gamma \) is apparently given by the solution of the Toeplitz system (2). From the viewpoint of robustness however, it is not recommended to use \( G^{-1} \) as a learning operator, due to modeling errors, noise and non-periodic disturbances. Moreover, \( G_{ILC}^{-1} \) is singular, because of its zero diagonal. In case of non-minimum phase systems \( G_{ILC}^{-1} \) will have further singularities. Since discrete time systems of order \( o > 2 \) are non-minimum phase [3], this is relevant for many ILC/RC applications.

III. 2-NORM OPTIMAL DESIGN

Optimal ILC/RC is based on the idea of minimizing an optimality criterion \( J_k \) in terms of the 2-norm. In order to find an optimal increment \( \Delta \mathbf{u}_k \) for each cycle \( k \), the criterion is minimized for each cycle \( k \). This is equivalent to gradient type methods [21], [5] and can be used for non-linear systems as well. Since the systems under consideration are linear and cycle invariant \( J_k = J \) will be the same for each cycle.

The optimality criterion \( J \) generally contains the weighted error \( \mathbf{e}_{k+1} \) and the weighted increment \( \Delta \mathbf{u}_k \) [11]. Sometimes the weighted actuating variable \( \mathbf{u}_{k+1} \) is also considered [4].

A. Optimal Learning Algorithm

The optimality criterion takes the form
\[
J = C_{k+1} Q C_{k+1}^T R C_{k+1} + \Delta \mathbf{u}_k^T R \Delta \mathbf{u}_k + \mathbf{u}_{k+1}^T P \mathbf{u}_{k+1}
\]

with symmetric and positive definite weight matrices \( Q, R \) and \( P \). The weight matrices determine the speed of learning with \( Q \), the change in actuating variable with \( R \) and the absolute values of the actuating variable with \( P \).

Setting the derivative of \( J \) with respect to \( \Delta \mathbf{u}_k \) to zero
\[
\frac{\delta J}{\delta \Delta \mathbf{u}_k} = -2 \mathbf{G}^T Q \mathbf{e}_k + 2 \mathbf{G}^T Q \mathbf{G} \Delta \mathbf{u}_k
\]

and solving for \( \Delta \mathbf{u}_k \) yields the optimal increment
\[
\Delta \mathbf{u}_k = (\mathbf{G}^T Q \mathbf{G} + \mathbf{R} + \mathbf{P})^{-1} (\mathbf{G}^T Q \mathbf{e}_k - \mathbf{P} \mathbf{u}_k).
\] (8)
The optimal learning matrix \( \Gamma \) and memory matrix \( S \) can thus be determined to
\[
\Gamma = (\mathbf{G}^T Q \mathbf{G} + \mathbf{R} + \mathbf{P})^{-1} \mathbf{G}^T Q
\]
\[
S = I - (\mathbf{G}^T Q \mathbf{G} + \mathbf{R} + \mathbf{P})^{-1} \mathbf{P}
\] (9)
by comparison of (8) with (6).

Since the design of \( \Gamma \) lies in the focus of this work, \( P \) will be neglected in the following. The influence of \( \mathbf{P} \) on \( \Gamma \) is the same as the influence of \( \mathbf{R} \) on \( \Gamma \). Therefore, \( \mathbf{P} \) can be considered in the choice of \( \mathbf{R} \). The rejection of \( \mathbf{P} \) yields the same \( \Gamma \) for structural and parameter optimization. In structural optimization \( J \) is minimized with respect to \( \Gamma \). In parameter optimization \( J \) is minimized with respect to \( \Delta \mathbf{u}_k \).

Setting \( \mathbf{P} = 0 \) in (9) yields \( S = I \), which is required for zero-error convergence [19]. As denoted above, \( \mathbf{P} \) weights the absolute values of \( \mathbf{u}_k \). The desired vector \( \mathbf{w} \) can only be reached if the corresponding \( \mathbf{u}_k \) is put on the system. Still, the choice of \( S \neq I \) can be required for certain applications. Note however, that \( S \) obtained from optimal design will always have full rank for positive definite weight \( P \). Stability statements in the following will thus not be influenced if \( P \) is neglected. For details regarding the function and design of \( S \) it is referred to [13].

Optimal design can be further simplified by replacing \( Q \) and \( \mathbf{R} \) with a single weight \( E \).

Consider \( Q = q I \) and \( \mathbf{R} = r I \) being weighted identity matrices as generally suggested [1], [6]. Then the learning operator can be written as
\[
\Gamma = (\mathbf{G}^T Q + q^{-1} R)^{-1} \mathbf{G}^T
\]
\[
= (\mathbf{G}^T Q + E)^{-1} \mathbf{G}^T
\] (10)
with weighted identity matrix \( \mathbf{E} = \frac{r}{q} \mathbf{I} \). It becomes clear that \( \mathbf{Q} \) will have no effect if \( \mathbf{R} = 0 \). Choice of neutral \( \mathbf{Q} = \mathbf{I} \) yields the single weight \( \mathbf{E} = \mathbf{R} \). The optimality criterion (7) becomes
\[
J = \mathbf{E}^\top \sum_{k=1}^{\infty} \Delta u_k^2 \mathbf{E} \Delta u_k + \Delta y_k^2 \mathbf{F} \Delta u_k
\]

B. Regularization and Generalized Inverse

Use of learning operator (10) is equivalent to solving the inverse problem of the Toeplitz system (2) using Tikhonov regularization.

As pointed out (2) does not have a solution for \( \mathbf{G} \). A generalized inverse of ill-conditioned \( \mathbf{G} \) will provide a good approximation. The limit
\[
\mathbf{G}^+ = \lim_{\eta \to 0} (\mathbf{G}^\top \mathbf{G} + \eta^2 \mathbf{I})^{-1} \mathbf{G}^\top
\]

defines the generalized inverse of real-valued \( \mathbf{G} \) [10]. This type of regularization is called Tikhonov regularization. Comparison of (11) with (10) reveals that the optimal design with choice of \( \mathbf{E} = \eta^2 \mathbf{I} \)
\[
(12)
\]
is equivalent to Tikhonov regularization.

Weight matrix \( \mathbf{E} \neq 0 \) does not only determine the convergence speed. It enables inversion of ill-conditioned \( \mathbf{G} \).

Consider the singular value decomposition of the system matrix
\[
\mathbf{G} = \mathbf{D} \Sigma \mathbf{V}^\top
\]

with singular values \( \Sigma = \text{diag}([\sigma_0, \ldots, \sigma_i, \ldots, \sigma_{N-1}]) \). The singular values equal the square root of the eigenvalues of
\[
\mathbf{G}^\top \mathbf{G} = \mathbf{V} \Sigma^2 \mathbf{V}^\top.
\]

The central question in optimal control design lies in the choice of weighting matrix \( \mathbf{E} \). To answer this question the stability in iteration domain \( k \) is analyzed.

Consider the MIMO control problem (6). Introduction of the complex shift variable \( z = e^{sT_k} \), in the iteration domain \( k \) analogous to the z-transform in the discrete time domain \( i \) yields the cyclic transfer function
\[
y'(z) = \mathbf{G} [z \mathbf{I} - \mathbf{S} + \mathbf{G} \mathbf{F}]^{-1} \mathbf{G} w(z)
\]

as presented in [19], [13]. Convergence can now be analyzed by the position of the poles \( z_{pi} \) with \( i \in [0, N-1] \) determined by
\[
\det(z_{pi} \mathbf{I} - \mathbf{S} + \mathbf{G} \mathbf{F}) = 0.
\]

Learning operator (10) inserted in (13) with \( \mathbf{S} = \mathbf{I} \) yields
\[
diag(z_{pi}) = \mathbf{V}^\top \left( \mathbf{I} - (\mathbf{G}^\top \mathbf{G} + \mathbf{E})^{-1} \mathbf{G}^\top \mathbf{G} \right) \mathbf{V}
\]

with unitary matrix \( \mathbf{V} \), which columns contain the eigenvectors \( \mathbf{v}_i \). The ILC/RC problem will converge stable if \( \left| z_{pi} \right| < 1, \forall i \) is assured.

A. Classical Choice of Weights

Generally the weighting matrix is chosen to be a positive weighted identity matrix [1], [6], [18]. This allows a direct translation of the weight \( \eta_i \) to the position of the poles. Since the singular values of the identity matrix are the elements of the identity matrix itself, (15) becomes
\[
diag(z_{pi}) = I - (\Sigma^2 \Sigma + \eta^2 \mathbf{I})^{-1} \Sigma^2 \Sigma.
\]

Optimal learning designs with weight (12) are stable, because
\[
z_{pi} = 1 - \frac{\sigma_i^2}{\sigma_i^2 + \eta^2} < 1, \quad \forall 0 < \eta < \infty
\]

holds. Large singular values yield poles that are placed close to zero. On the contrary small singular values yield poles close to 1 for the same choice of \( \eta^2 \).

Under assumption of an accurate model, optimal design generates a distribution of real-valued poles between 0 and 1 and will always be stable. Convergence will take place monotonically with different speeds for each eigenvector \( \mathbf{v}_i \). Very small \( \eta^2 \) will tend to one step convergence, as all poles are placed close to zero.

The eigenvectors \( \mathbf{v}_i \) corresponding to each pole \( z_{pi} \) determine the time behavior of the ILC/RC within the cycle. With the presented choice of \( \mathbf{E} \) optimal design does not enable to influence \( \mathbf{V} \). Only the position of the poles can be determined by choice of \( \eta^2 \) in correspondence to the singular values of \( \mathbf{G} \).

B. Sophisticated Choice of Weights

The choice of more complex weights might not only place the poles, but also determine the eigenvectors \( \mathbf{v}_i \). This is generally not desired. Individually placing each pole while keeping the \( \mathbf{V}_i \) determined by \( \mathbf{G}^+ \) is more desirable.

In order to weight certain time values \( i \) of the signal \( u_k(i) \) it could be considered to choose an individually weighted diagonal matrix
\[
\mathbf{E} = \text{diag} [\eta_0^2, \ldots, \eta_i^2, \ldots, \eta_{N-1}^2].
\]

Considering (15) reveals what happens if a weighted diagonal weighting matrix is used. The eigenvalues of the weighting matrix (18) are not equivalent to the eigenvector matrix of \( \mathbf{G}^+ \). Therefore, pole placement cannot be described analogous to (16). A simple relation between pole placement and choice of weights cannot be derived.
Individually placing each pole while keeping up the eigenvectors \( \mathcal{E} \) can be achieved by calculating the weight matrix in reverse

\[
\hat{E} = V E V^* \tag{19}
\]

with \( \hat{E} = \text{diag} \left[ \eta_0^2, \cdots, \eta_i^2, \eta_{N-1}^2 \right] \) from the pole placement equation

\[
z_{pi} = 1 - \frac{\sigma_i^2}{\sigma_i^2 + \eta_i^2}. \tag{20}
\]

C. Interpretation of the Eigenvectors

The eigenvectors \( \mathcal{E}_i \) of \( G^T G \) have remarkable characteristics, because system matrix \( G_{ILC} \) in (4) and cyclic matrix \( G_{RC} \) in (5) are Toeplitz matrices. For the cyclic system matrix (5) each \( \mathcal{E}_i \) corresponds to a single frequency contained in the cyclic signal. The singular values correspond to the magnitude of each frequency component. It has been shown for \( G_{ILC} \) that the same holds true in good approximation [13].

The described optimal designs can therefore be interpreted as weighting each frequency component corresponding to its magnitude determined by (17) or (20). By (17) higher frequencies of low-pass systems will be attached with low learning gains. This is usually accommodating, since the ILC/RC will be more robust in regard to modeling inaccuracies for high frequencies and noise. By (20) the poles can be placed with regards to other system specific characteristics as e. g. band limited disturbances.

However, note that \( z_{pi} \not= 1 \) holds \( \forall i \), because of \( \eta_i \not= 0 \) \( \forall i \). Thus, every eigenvector \( \mathcal{E}_i \) will eventually be considered in learning, even if the corresponding gain is very small. This limits robustness of optimal designs. In order to have stable poles for the real system, the nominal poles have to be placed in between

\[
1 - 2 \cos (|\Delta \varphi_i|) < |z_{pi}| < 1 \tag{21}
\]

with \( |\Delta \varphi_i| < 90^\circ \) and \( \Delta \varphi_i = \varphi_{\text{real}} - \varphi_i \) being the difference from real to nominal phase of the plant [13]. The allowed modeling error is thus limited to \( |\Delta \varphi_i| < 90^\circ \) for all frequencies.

D. Comparison to other ILC designs

Optimal learning falls into the group of phase canceling learning designs. As the matrix \( G^T \hat{G} + \hat{E} \) is symmetrical for all of the above choices of \( \hat{E} \) it does not contain any phase information. The entire phase information of the learning operator (10) is contained in \( G^T \).

A classical phase canceling design is the contraction mapping or gradient based learning law [7], [14]

\[
\Gamma = \gamma G^T. \tag{22}
\]

The only tuning parameter of this law is \( \gamma \), which determines the step length for learning from cycle to cycle. The limit

\[
\lim_{\eta \to \infty} (G^T G + \eta^2 \hat{L})^{-1} G^T = \frac{1}{\eta^2} G^T \tag{23}
\]

shows that contraction mapping can be considered as a special case of classic optimal control as in IV-A. For a small stationary gain system with large choice of \( \eta^2 \gg \sum_{i=0}^{N-1} g_i \)

the similarity of optimal design to contraction mapping increases. Admittedly, an optimal design similar to contraction mapping tends to a defensive choice of learning gain \( \gamma \) as limit (22) reveals. However, the advantage of optimal design over contraction mapping is the automatic choice of learning gain. This assures stable convergence as shown by (17), which cannot be guaranteed for contraction mapping. E. g. for systems with stationary gain larger one the weight \( \gamma \) becomes very important for contraction mapping to obtain stability.

Another, more aggressive phase canceling design is partial isometry learning [14]

\[
\hat{E} = \hat{S} = \sum^+ \sum^T \hat{S} \tag{24}
\]

with \( \sum^+ \) being the pseudo-inverse of \( \sum \). Partial isometry shows analogies to the more sophisticated type of optimal learning design as in IV-B. Consider the choice of

\[
\gamma \Gamma = \hat{L} = \sum^T \sum^+ \hat{S} \tag{25}
\]

in (19). Both learning designs are equivalent as

\[
\lim_{\alpha \to 0} \left( \Gamma^T \Gamma + \alpha \hat{E} \right)^{-1} \Gamma^T = \sum^T \sum^+ \hat{S} \tag{26}
\]

reveals. The limit is necessary in case of any \( \sigma_i = 0 \). As stated above, this will always be the case for matrix (4). Partial isometry is less conservative than contraction mapping, because it includes a partial inversion of the magnitude of the system. Analogous to contraction mapping parameter \( \gamma \) becomes very important for systems with stationary gain larger than one.

V. REDUCED IMPLEMENTATION

Optimal designs are well suited for a reduced implementation of ILC and RC.

Consider the calculation of a single time value \( u_{k+1}(i) \) of the actuating vector \( u_{k+1} \) in (6)

\[
u_{k+1}(i) = u_k(i) + \gamma_i e_k
\]

with \( \gamma_i \) being the \( i \)-th row of \( \Gamma \). Further, consider the continuous application of \( \gamma_i \) as the FIR-filter \( \gamma_i \) with transfer function

\[
\gamma_i(z) = \sum_{l=0}^{N-1} \gamma_{i,l} z^{-l} \tag{27}
\]

on the error signal \( e \). Defining the cyclic shift \( z^{-N} \) with

\[
u_{k+1} = u_k z^{-N} \tag{28}
\]
the update (23) can be written as the internal model transfer function
\[
\frac{u(z)}{e(z)} = \frac{\gamma_i(z)z^{-(N-i)}}{1 - z^{-(N)}}
\] (26)
as presented in [13].

The FIR-filter \( \gamma_i \) of a middle time step of the cycle, e. g. \( i = N/2 \), obtained from optimal designs will generally decay to zero on both sides as depict in fig. 2. Recalling the interpretation of singular values as the magnitude of the frequency spectrum, this decay is a result of the smooth attenuation of higher frequency components inherent in classical optimal design for low-pass systems. This corresponds to a smooth windowing in digital filter design, which leads to small leakage effects in the time signal. No leakage provides unaltered frequency response of the FIR-filter regardless to the input signal. [16]

Therefore, it is suggested to design the optimal learning matrix \( \Gamma \) on a sampled interval as short as possible with regards to the length \( N \) of the impulse response of the system. Then, take a middle row \( i = N/2 \) of the optimal \( \Gamma \) to obtain the FIR-filter \( \gamma_{N/2} \). The filter \( \gamma_{N/2} \) is then applied on the cycle of length \( N + M \) yielding
\[
\frac{u(z)}{e(z)} = \frac{\gamma_i(z)z^{-(N-i+M)}}{1 - z^{-(N+M)}} \quad \text{with } i = N/2.
\] (27)
The learning characteristics of the reduced implementation (27) are identical to the full cycle implementation (26) if no leakage exists, which is denoted, is nearly the case for classical optimal design on a low-pass system. The calculations for each time step are reduced from \( N + M \) to \( N \).

Note that the internal model (26) is an exact infinite time system description if the cycles follow each other consecutively as in RC. It is an approximation when applied on a finite time length \( T_c \) as in ILC, because by (25) a change-over from cycle to cycle is included to the model. The system model (26) is a Fourier basis model. The boundary effects of Fourier basis model designs for application in ILC have been elaborated comprehensively in [13].

VI. VALIDATION ON A TESTBED

Two optimal RC designs are presented in application on a testbed.

The testbed is an AC servo drive positioning a swivel, which is connected to the drive by a torsion rod. The swivel repetitively follows a desired input
\[
u(i) = A/2 + A \sin(2\pi f i T_s + \varphi)
\]
with \( A = 0.2, \varphi = -90^\circ, f = 1 \) Hz. The discrete time transfer function with sampling time \( T_s = 1 \) ms is \( G(z) = \frac{7.366e^{-7z^3} + 8.002e^{-6z^2} + 7.032e^{-6z} + 7.173e^{-7}}{z^4 - 3.891z^3 + 5.739z^2 - 3.806z + 0.9567} \). The discrete system is non-minimum phase with zeros
\[z_1 = -9.7719, \quad z_2 = -0.9912, \quad z_3 = 0.1005.\]
The system has two resonance frequencies at
\[f_1 \approx 2.85 \text{ Hz,} \quad f_2 \approx 41 \text{ Hz.}\]

Both controller are obtained by use of (10) with the classical choice of weight as in (12). Parameter \( \eta \) is chosen to be \( \eta^2 = 0.01 \) and the learning matrix \( \gamma^0 \) is weighted with \( \gamma = 0.6 \). The design length is chosen to be \( N = 500 \) corresponding to \( T = 0.5 \) s, which approximately fits the settling time of the impulse response. The cycle length is \( T_c = 1 \) s, which matches the period of desired signal \( u \).

The first controller is designed on a reduced order model neglecting the higher resonance \( f_r \),
\[G_{\text{red}}(z) = \frac{0.0001336z + 0.0001326}{(z)^2 - 1.976(z) + .9766}.
\]
The second controller is designed on the full order model.

Fig. 1 depicts the frequency spectra of the full order and reduced order model and the frequency spectra of the corresponding FIR-filters \( \gamma_{N/2} \). It becomes illustrative that optimal design belongs to the group of phase canceling learning designs. For both designs the phase information is inverted exactly. The magnitude equates to a system inversion up to approximately \( f = 7 \) Hz. Higher frequencies are damped according to the frequency spectrum of the system model.

![Fig. 1. Frequency spectrum of the full order and reduced order model and the corresponding frequency spectra of the FIR-learning filters \( \gamma_{N/2} \).](image-url)
similar error within the first 8 cycles. Controller 1, designed on the reduced model, seems to converge stable after the first cycles. The system eventually diverges, because the robustness restrictions $|\varphi_i| < 90^\circ$ from (21) are not fulfilled for high frequency components as illustrated in Fig. 1. Long term stability is only provided by controller 2, which adequately accounts higher frequency components. Controller 2 also practically proofs the suitability of optimal design for non-minimum phase systems.

VII. CONCLUSION

2-norm optimal learning control design has been studied. It has been shown, that optimal designs are well suited for ill-conditioned and non-minimum phase systems. On the other hand, optimal design necessitates good system knowledge, because stability will be guaranteed for model errors in phase in the range of $-90^\circ < \Delta \varphi_i < 90^\circ$ over all frequencies only. A frequency interpretation of the learning weights has been suggested. This allows a decided choice of weights and a sophisticated weighting that can e. g. be used to suppress band limited disturbances. Also, a novel implementation of ILC/RC has been proposed that enables to utilize all features of classical optimal controller design without the drawback of handling large matrices for long cycles.

Fig. 2. Structure of optimal learning matrix $\Gamma$.

Fig. 3. Mean square error of classical optimal RC with reduced implementation, $\gamma = 0.6$, $\eta^2 = 0.01$.

REFERENCES