01 Jan 2011

Towards Transient Growth Analysis and Design in Iterative Learning Control

Douglas A. Bristow
Missouri University of Science and Technology, dbristow@mst.edu

John R. Singler
Missouri University of Science and Technology, singlerj@mst.edu

Recommended Citation

D. A. Bristow and J. R. Singler, "Towards Transient Growth Analysis and Design in Iterative Learning Control," *International Journal of Control*, Taylor & Francis, Jan 2011.
The definitive version is available at https://doi.org/10.1080/00207179.2011.596224

Towards Transient Growth Analysis and Design in Iterative Learning Control

Douglas A. Bristow[1] and John R. Singler[2]

[1] *Mechanical and Aerospace Engineering, Missouri University of Science and Technology, Rolla, MO, U.S.A.*
[2]*Mathematics and Statistics, Missouri University of Science and Technology, Rolla, MO, U.S.A.*

Corresponding author: Douglas A. Bristow, Dept. of Mech. and Aero. Eng., Missouri University of Science and Technology, 292C Toomey Hall, Rolla, MO 65409, ph: 573-341-6559, dbristow@mst.edu

In this paper the problem of bounding transient growth in Iterative Learning Control (ILC) is examined. While transient growth is not a desirable property, the alternative, robust monotonic convergence, leads to fundamental performance limitations. To circumvent these limitations, this paper considers the possibility that some transient growth, if properly limited, is a viable and practical option. Toward this end, this paper proposes tools for analyzing worst-case transient growth in ILC. The proposed tools are based on pseudospectra analysis, which is extended to apply to ILC of uncertain systems. Two practical problems in norm-optimal ILC weighting parameter design are considered. Using the presented tools, it is demonstrated that successful design in the transient growth regime is possible, i.e. the transient growth is kept small while significantly improving asymptotic performance, despite model uncertainty.

Keywords: iterative learning control, pseudospectra, transient analysis, robustness

I. Introduction

Iterative learning control (ILC) [1-3] is used to improve the performance of systems that repeat the same operation many times. At the conclusion of each operation, the tracking error is used to batch-update a feedforward control signal. Convergence of the learning process results in a feedforward control signal that is customized for the repeated motion, yielding very low tracking error.

ILC is a performance-improving control algorithm, rather than a stabilizing algorithm, and thus the emphasis of much of the ILC literature focuses on behavior at convergence. Although convergence of the algorithm is typically demonstrated, comparatively little attention is given to the nature of the convergence. The transient behavior of the learning process, however, is critically important in many practical

applications. For example, in robotics and manufacturing applications, slow convergence leads to delays in process startup and possibly costly material waste. Perhaps of greater concern to the ILC designer is the problem of large transient growth [4], whereby the error may grow rapidly and with little warning, potentially damaging hardware. Generally, when large transient growth occurs the learning process must be aborted, and thus the high performance guaranteed at convergence is never realized.

The problem of large transient growth has been studied extensively by Longman and colleagues [4-8]. In many cases the ILC signal appears to be converging over early iterations, followed by rapidly diverging behavior [4]. The rapid divergence may be the onset of a large transient growth, or it may be true instability. In practice the two possibilities can be indistinguishable. Further complicating the problem is the fact that large transient growth can occur in exponentially stable learning systems [5]. Essentially, the problem derives from the two-dimensional, time- and iteration-domain, dynamics of the learning process (in this context, the dimension refers to the domain over which the dynamics evolve, not the number of state variables in the dynamic system). Generally, more stringent stability conditions are desired in two-dimensional systems [9].

A major thrust in ILC research has been the development of algorithms that ensure contraction of the error, measured by an appropriate norm, from one iteration to the next, termed monotonic convergence. Popular design techniques for achieving monotonic convergence are frequency domain design [5], norm-optimal design [10], model-inversion design [11], and gradient design [12]. When the plant model is known precisely, these methods are capable of converging arbitrarily quickly to an arbitrarily low error (in no-noise scenarios). However, in practice the plant model is

never precisely known. An algorithm that yields monotonic convergence for a bounded set of plant models, containing the actual plant, is said to be robustly monotonically convergent (RMC). The above designs can be used to construct RMC algorithms, but converged tracking performance is limited, depending on the size of the uncertainty bounds [11-14].

Due to the limited performance at convergence, RMC is often overly restrictive and generally conflicts with the high performance goals of ILC. Specifically, RMC algorithms are very conservative in that the worst case system is forced to convergence monotonically; thus, RMC sacrifices performance for a restrictive monotonicity guarantee. However, monotonic convergence is not a necessary constraint in many applications. Therefore, some transient growth, if kept small, is acceptable for many applications, particularly if the result is a significant performance increase.

In order to extend ILC design beyond the limitations of RMC, it is necessary to revisit the problem of transient growth. The main contribution of this work is the application of pseudospectra analysis [15] to the ILC transient growth problem. In particular, pseudospectra tools are extended to examine worst-case transient growth in ILC systems with model uncertainty. The use of pseudospectra tools for ILC transient analysis was first presented in [16], although that work did not consider model uncertainty. The approach presented in [16] and this paper are the first such approaches that allows the ILC designer to analyze the effect of algorithm or parameter changes on the transient learning behavior. Although it may be possible to use these tools to generate fundamentally new ILC algorithms, that topic is not explored here. Instead, application of the proposed tools is illustrated by examining two parameter design problems in norm-optimal ILC. Using the presented tools,

successful design in the transient growth regime is demonstrated, i.e. transient growth is limited while significantly improving asymptotic performance, despite model uncertainty.

Throughout this work let $\rho(\cdot)$ denote the spectral radius and let $\|\cdot\|$ denote the Euclidean 2-norm on $\mathbb{C}^n$ and also the induced matrix norm on $\mathbb{C}^{nxm}$, that is for a matrix $\mathbf{T} \in \mathbb{C}^{nxm}$,

$$\|\mathbf{T}\| = \sup_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}\in\mathbb{C}^m}} \|\mathbf{T}\mathbf{x}\| = \bar{\sigma}(\mathbf{T}) \text{ and } \rho(\mathbf{T}) = \max|\lambda(\mathbf{T})|,$$

where $\bar{\sigma}(\cdot)$ is the maximum singular value and $\lambda(\cdot)$ is the set of eigenvalues.

The remainder of this paper is organized as follows. In Section II key results in monotonic convergence analysis for ILC are reviewed. The results illustrate key performance limitations that are imposed by the RMC condition. There are a number of mathematical tools that can be applied to the analysis of non-monotonic learning; these are surveyed in Section III. One tool, the pseudospectra, is selected for further exploration, and in Section IV methods are developed for analyzing worst-case transient growth for a set of uncertain systems. Section V demonstrates the application of this tool on two norm-optimal ILC design problems: convergence rate weighting and time-varying Q-filter shaping. Finally, conclusions are given in Section VI.

II. Background: Stability and Monotonic Convergence in ILC

Consider a discrete-time, single-output linear time-invariant system,

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + d(k), \end{aligned}$$

where $u(k)$ is the input and $y(k)$ is the output. The output response to external disturbances and the initial conditions are lumped in $d(k)$. The system can be "lifted" [17] by noting that over the finite time horizon, $k = m, m+1, \ldots, m+N-1$,

$$\underbrace{\begin{bmatrix} y(m) \\ y(m+1) \\ \vdots \\ y(m+N-1) \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} p_m & 0 & 0 & \cdots & 0 \\ p_{m+1} & p_m & 0 & \ddots & \vdots \\ p_{m+2} & p_{m+1} & p_m & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ p_{m+N-1} & \cdots & p_{m+2} & p_{m+1} & p_m \end{bmatrix}}_{\mathbf{P}} \underbrace{\begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(N-1) \end{bmatrix}}_{\mathbf{u}} + \underbrace{\begin{bmatrix} d(m) \\ d(m+1) \\ \vdots \\ d(m+N-1) \end{bmatrix}}_{\mathbf{d}}, \quad (1)$$

where,

$$p_i = \begin{cases} 0, & i < m \\ CA^{i-1}B, & i \geq m \end{cases}$$

are the system Markov parameters and m is the relative degree. Let the desired output be given by $y_d(k)$, or

$$\mathbf{y_d} = \begin{bmatrix} y_d(m) & y_d(m+1) & \cdots & y_d(m+N-1) \end{bmatrix}^T.$$

In the ILC problem it is desired that the system repeat a process indefinitely. Each iteration, $j = 0,1,\ldots$ of the process involves tracking y_d, then resetting to zero initial conditions. It is standard to assume that the disturbance repeats each iteration, $\mathbf{d} = \mathbf{d}_0 = \mathbf{d}_1 = \cdots$. Then, the tracking error on the j^{th} iteration is given by $\mathbf{e}_j = \mathbf{y_d} - \mathbf{y}_j$, or,

$$\mathbf{e}_j = -\mathbf{P}\mathbf{u}_j + \mathbf{e}_0, \; j = 1,2,\ldots. \quad (2)$$

The ILC problem is to select $\mathbf{u}_j$ using the error history from previous trials to asymptotically reduce the error. A common method is the first-order linear ILC update algorithm, given by,

$$\mathbf{u}_{j+1} = \mathbf{Q}\left(\mathbf{u}_j + \mathbf{L}\,\mathbf{e}_j\right), \quad (3)$$

where Θ and Λ are in R^{NxN}. Combining (1), (3), closed-loop dynamics are given by

$$\mathbf{u}_{j+1} = \mathbf{T}\mathbf{u}_j + \mathbf{f}_0, \quad (4)$$

where $\mathbf{T} = \mathbf{Q}\left(\mathbf{I} - \mathbf{L}\,\mathbf{P}\right)$ and $\mathbf{f}_0 = \mathbf{Q}\,\mathbf{L}\,\mathbf{e}_0$. Clearly the ILC system is convergent if $\rho(\mathbf{T}) < 1$.

Remark: In some design techniques, such as frequency domain [5] or LMI [18], the update algorithm is written with dynamic filters. These algorithms can be written as (3) by lifting the filters into matrices $\boldsymbol{\Theta}$ and $\boldsymbol{\Lambda}$. However, in other ILC approaches the algorithm is feedback-based, such as [19], or higher-order. The analysis presented in this work applies only to algorithms of the form (3).

Remark: Contrary to many results in the field of control systems, convergence, alone, is often insufficient in ILC because of the propensity for slow convergence and large transient growth [5,6]. For applications where large transient growth and slow convergence may not be problematic, it is well known that $\mathbf{Q} = \mathbf{I}$ and $\mathbf{L} = c\mathbf{I}$, where c has the same sign and smaller magnitude than p_m, is sufficient for zero error convergence [5].

To analyze the transient behavior of an exponentially convergent ILC system, define $\mathbf{u}_\infty \triangleq \lim_{j\to\infty} \mathbf{u}_j$ and $\delta_j\mathbf{u} \triangleq \mathbf{u}_\infty - \mathbf{u}_j$. Then, from (4),

$$\delta_{j+1}\mathbf{u} = \mathbf{T}\delta_j\mathbf{u}, \tag{5}$$

or,

$$\delta_{j+1}\mathbf{u} = \mathbf{T}^j\delta_0\mathbf{u}. \tag{6}$$

Therefore, the sequence

$$\|\mathbf{T}\|, \|\mathbf{T}^2\|, \ldots, \|\mathbf{T}^j\|, \ldots \tag{7}$$

gives the worst-case bound on the overall transient growth or decay of $\|\delta_j\mathbf{u}\|$ during learning. The special case $\|\mathbf{T}\| < 1$ is referred to as *monotonic convergence* because the bounding sequence is necessarily monotonically decreasing, $\|\mathbf{T}^{j+1}\| \le \|\mathbf{T}\|\|\mathbf{T}^j\| < \|\mathbf{T}^j\|$.

III. The Pseudospectra

For a given $N \times N$ matrix $\mathbf{T}$, the key mathematical problem in this work is analyzing the behavior of (7). If a matrix $\mathbf{T}$ is *normal*, i.e., $\mathbf{T}\mathbf{T}^* = \mathbf{T}^*\mathbf{T}$, then (7) is

trivially given by $\|\mathbf{T}^k\| = \rho(\mathbf{T})^k$ for all k. More generally, tight bounds on (7) are nontrivial; some well known bounds are summarized in Table 1 and illustrated in Figure 1.

Table 1. Well-known transient response bounds.

Spectral radius decay rate[1]:	$\|\mathbf{T}^k\| \le \kappa(\mathbf{V})\rho(\mathbf{T})^k$
Norm bound:	$\|\mathbf{T}^k\| \le \|\mathbf{T}\|^k$
Growth rate:	$\|\mathbf{T}^{k+1}\| / \|\mathbf{T}^k\| \le \|\mathbf{T}\|$
Limiting rate[2]:	$\lim_{k\to\infty} \frac{1}{k} \log\|\mathbf{T}^k\| = \rho(\mathbf{T})$

[1] $\kappa(\mathbf{V}) = \|\mathbf{V}\|\|\mathbf{V}^{-1}\|$, Proof on page 19, [15].
[2] Proof on page 159, [15].

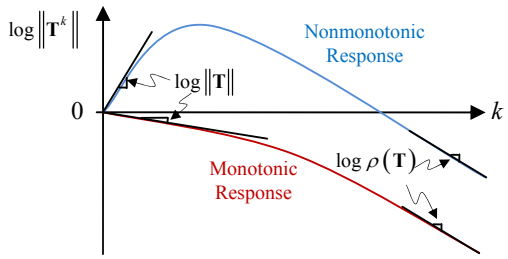

Figure 1. Monotonic and nonmonotonic possible transient behavior of $\|\mathbf{T}^k\|$.

Recent research has focused on bounding (7) for nonnormal matrices. Three approaches are departure from normality [21-23], Lyapunov inequalities [24,25], and pseudospectra [15]. Here the pseudospectra is used because it has a history of success in engineering applications and a visual interpretation useful in design. Specifically, the pseudospectra and associated pseudospectral radius are defined [15] as follows:

Definition: The ε-pseudospectra of a matrix $\mathbf{T}$ is the set $\sigma_\varepsilon(\mathbf{T})$ in the complex plane consisting of all points $z \in \mathbb{C}$ such that z is an eigenvalue of $\mathbf{T}+\mathbf{E}$ for some $\mathbf{E} \in \mathbb{C}^{n\times n}$ with $\|\mathbf{E}\| < \varepsilon$. The ε-pseudospectral radius of $\mathbf{T}$ is defined by

$$\rho_\varepsilon(\mathbf{T}) = \{\max|z| : z \in \sigma_\varepsilon(\mathbf{T})\}.$$

Equivalently, the pseudospectra is the set where the resolvent $(z\mathbf{I}-\mathbf{T})^{-1}$ is large:

$$\sigma_\varepsilon(\mathbf{T}) = \{z \in \mathbb{C} : \|(z\mathbf{I}-\mathbf{T})^{-1}\| > \varepsilon^{-1}\}.$$

The eigenvalues of **T** are the points where the norm of the resolvent is infinite, therefore, $\lambda(\mathbf{T}) \subset \sigma_\varepsilon(\mathbf{T})$ for all $\varepsilon > 0$.

Many pseudospectral bounds for transient growth can be found in [15]. A key result is the Kreiss matrix theorem, which bounds the maximum transient growth in terms of the ε-pseudospectral radius:

Kreiss Matrix Theorem [15, Section 18]: For any $N \times N$ matrix **T,**

$$\mathrm{K}(\mathbf{T}) \le \sup_{k\ge 0} \|\mathbf{T}^k\| \le eN\mathrm{K}(\mathbf{T}),$$

where e is the exponential constant and the *Kreiss constant* is defined by

$$\mathrm{K}(\mathbf{T}) = \sup_{\varepsilon>0} \frac{\rho_\varepsilon(\mathbf{T})-1}{\varepsilon}.$$

The Kreiss constant gives upper and lower bounds on the maximum transient growth, however the constant is difficult to compute. Therefore, plots of pseudospectra are often used to estimate the magnitude of transient growth.

To illustrate the visualization aspect of the pseudospectra, a simple example is used. Consider the three matrices

$$\mathbf{T}_1 = \begin{bmatrix} 0.8 & 0 \\ 0.1 & 0.8 \end{bmatrix}, \mathbf{T}_2 = \begin{bmatrix} 0.8 & 0 \\ 1 & 0.8 \end{bmatrix}, \mathbf{T}_3 = \begin{bmatrix} 0.8 & 0 \\ 10 & 0.8 \end{bmatrix}.$$

Shown in Figure 2 are boundaries (or level sets) of pseudospectra for these matrices for various values of ε computed using Eigtool [30]. The eigenvalues are located at 0.8 for each system, however the pseudospectra of each matrix are quite different. For $\mathbf{T}_1$, the pseudospectra boundaries are clustered closely around its eigenvalues; for $\mathbf{T}_2$, the pseudospectra boundaries are larger and extend outside of the unit circle for $\varepsilon = 10^{-1}$; for $\mathbf{T}_3$, the pseudospectra extend well outside of the unit circle, even for very small values of ε.

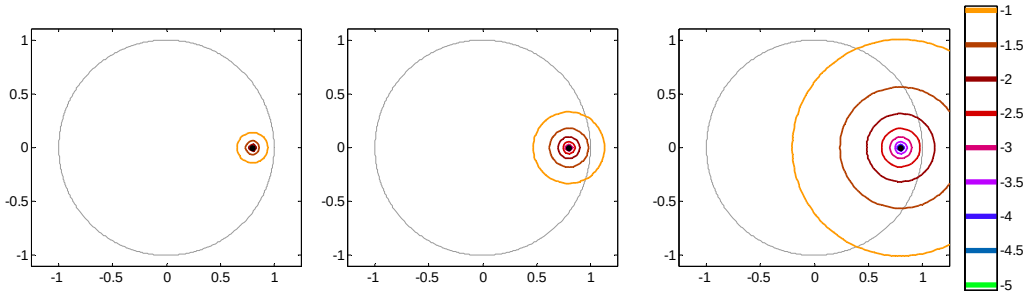

Figure 2. The ε-pseudospectra boundaries of $\mathbf{T}_1$, $\mathbf{T}_2$, and $\mathbf{T}_3$ from left to right. The color bar is on a $\log_{10}$ scale so that the values of ε are $10^{-5},\ldots,10^{-1}$ from inside to outside.

Interpretation of the pseudospectra plots is done as follows. The Kreiss matrix theorem shows that if the pseudospectral radius $\rho_\varepsilon(\mathbf{T})$ sufficiently exceeds one, then transient growth should be expected; more precisely, if $\rho_\varepsilon(\mathbf{T}) > 1+\varepsilon$ for some ε, then $\sup_{k\geq 0}\|\mathbf{T}^k\| > 1$ and therefore transient growth must occur. If the pseudospectral radius exceeds one for a small value of ε, large transient growth is expected.

Thus, Figure 2 above indicates that transient growth is expected for $\mathbf{T}_2$ and $\mathbf{T}_3$ with more transient growth expected for $\mathbf{T}_3$. This is confirmed by direct computations of the transients $\|\mathbf{T}_1^k\|$, $\|\mathbf{T}_2^k\|$, and $\|\mathbf{T}_3^k\|$, in Figure 3.

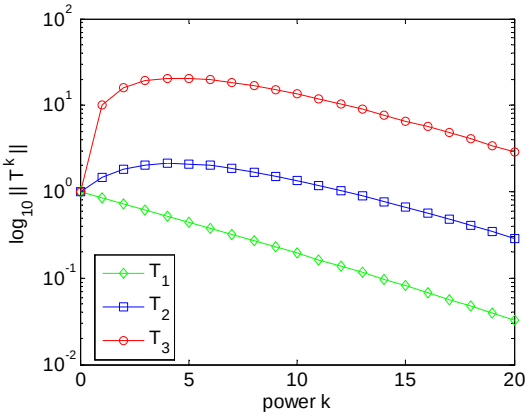

Figure 3. Transient behavior of the powers $\|\mathbf{T}_i^k\|$ for $i = 1,2,3$. For $i = 1$, the powers decay monotonically to zero; for $i = 2$, there is transient growth before decay; for $i = 3$, there is larger transient growth. These results are easily predicted using the pseudospectra plots above in Figure 2.

IV. Accounting for Uncertainty in Transient Growth Analysis

The previous transient growth analysis assumes the transition matrix is known exactly. However, in practice the system is not precisely known. Recall that the transition matrix is given by, $\mathbf{T} = \mathbf{Q}(\mathbf{I} - \mathbf{L}\mathbf{P})$, where $\mathbf{P}$ is the plant matrix. As is standard, [13,14], assume the plant $\mathbf{P}$ is uncertain but can be represented by

$$\mathbf{P} = \mathbf{P}_\mathbf{m}(\mathbf{I} + \mathbf{W}\boldsymbol{\Delta}) \tag{8}$$

where $\mathbf{P}_\mathbf{m}$ and $\mathbf{W}$ are known and $\boldsymbol{\Delta}$ is unknown. $\mathbf{P}_\mathbf{m}$, $\mathbf{W}$, and $\boldsymbol{\Delta}$, are lower triangular Toeplitz. With this uncertainty representation, the transition matrix $\mathbf{T}$ can be written as

$$\mathbf{T} = \mathbf{X} + \mathbf{Y}\boldsymbol{\Delta}, \tag{9}$$

where,

$$\begin{aligned} \mathbf{X} &= \mathbf{Q}(\mathbf{I} - \mathbf{L}\mathbf{P}_\mathbf{m}), \\ \mathbf{Y} &= -\mathbf{Q}\mathbf{L}\mathbf{P}_\mathbf{m}\mathbf{W}. \end{aligned} \tag{10}$$

Thus, the transient behavior of the powers of $\mathbf{T} = \mathbf{X} + \mathbf{Y}\boldsymbol{\Delta}$ are of interest, where $\mathbf{X}$ and $\mathbf{Y}$ are known matrices, but $\boldsymbol{\Delta}$ is an unknown lower triangular Toeplitz matrix.

Assuming a known bound on the uncertainty, say $\|\boldsymbol{\Delta}\| < \delta$, consider the following modified pseudospectra and pseudospectral radius:

Definition: For given matrices $\mathbf{X}, \mathbf{Y} \in \mathbf{R}^{NxN}$, define the modified pseudospectra $\sigma_{\delta,\varepsilon}(\mathbf{T})$ to be the set in the complex plane consisting of all points $z \in \mathbb{C}$ such that z is an eigenvalue of $\mathbf{T} = \mathbf{X} + \mathbf{Y}\boldsymbol{\Delta} + \mathbf{E}$ for some $\boldsymbol{\Delta} \in \mathbf{R}^{N\times N}$ and $\mathbf{E} \in \mathbb{C}^{N\times N}$ with $\boldsymbol{\Delta}$ lower triangular Toeplitz and $\|\boldsymbol{\Delta}\| < \delta, \|\mathbf{E}\| < \varepsilon$. The modified pseudospectral radius is defined by

$$\rho_{\delta,\varepsilon}(\mathbf{T}) = \{\max|z| : z \in \sigma_{\delta,\varepsilon}(\mathbf{T})\}.$$

This modified pseudospectra yields a worst-case transient analysis. If $\sigma_{\delta,\varepsilon}(\mathbf{T})$ extends far enough outside of the unit disk, then there is a $\boldsymbol{\Delta}$ so that the powers of

$\mathbf{T} = \mathbf{X} + \mathbf{Y}\boldsymbol{\Delta}$ experience transient growth before decaying to zero. This set is difficult to compute directly, so two approaches are considered for approximating $\sigma_{\delta,\varepsilon}(\mathbf{T})$. In the first approach, an upper bounding method using spectral value sets is examined; in the second approach, a random uncertainty lower bounding method is used.

A. Approach 1: Spectral Value Sets

The following definition and theorem for spectral value sets will be useful in the derivation of an upper bounding set that contains the modified pseudospectra $\sigma_{\delta,\varepsilon}(\mathbf{T})$.

Definition: For given matrices $\mathbf{A} \in \mathbf{R}^{NxN}$, $\mathbf{B} \in \mathbf{R}^{NxM}$, and $\mathbf{C} \in \mathbf{R}^{LxN}$, the ε-spectral value set $\sigma_\varepsilon(\mathbf{A},\mathbf{B},\mathbf{C})$ is the set in the complex plane consisting of all points $z \in \mathbb{C}$ such that z is an eigenvalue of $\mathbf{A}+\mathbf{B}\boldsymbol{\Theta}\mathbf{C}$ for some $\boldsymbol{\Theta} \in \mathbb{C}^{M\times L}$ with $\|\boldsymbol{\Theta}\| < \varepsilon$.

Theorem 1 [24,26]: The set $\sigma_\varepsilon(\mathbf{A},\mathbf{B},\mathbf{C})$ is given by

$$\sigma_\varepsilon(\mathbf{A},\mathbf{B},\mathbf{C}) = \sigma(\mathbf{A}) \bigcup \left\{ z \in \mathbf{C}: \left\| \mathbf{C}(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B} \right\| > \varepsilon^{-1} \right\}.$$

The pseudospectra can be considered a special case of spectral value sets since $\sigma_\varepsilon(\mathbf{A}) = \sigma_\varepsilon(\mathbf{A},\mathbf{I},\mathbf{I})$. Now, define the set

$$\tilde{\sigma}_{\delta,\varepsilon}(\mathbf{T}) = \left\{ z \in \mathrm{C} : \left\| \begin{bmatrix} \mathbf{I} \\ \mathbf{I} \end{bmatrix} (z\mathbf{I}-\mathbf{X})^{-1} \begin{bmatrix} \frac{\delta}{\varepsilon}\mathbf{Y} & \mathbf{I} \end{bmatrix} \right\| > \varepsilon^{-1} \text{ or } z \in \sigma(\mathbf{X}) \right\}, \qquad (11)$$

which is a larger bounding set for the modified pseudospectra, as shown in the following theorem.

Theorem 2: Let $\mathbf{T}$ be given by (9), and $\tilde{\sigma}_{\delta,\varepsilon}(\mathbf{T})$ given by (11). Then $\sigma_{\delta,\varepsilon}(\mathbf{T}) \subset \tilde{\sigma}_{\delta,\varepsilon}(\mathbf{T})$.

Proof: Assume $z \in \sigma_{\delta,\varepsilon}(\mathbf{T})$. Then, z is an eigenvalue of $\mathbf{T}+\mathbf{E} = \mathbf{X}+\mathbf{Y}\boldsymbol{\Delta}+\mathbf{E}$, with $\boldsymbol{\Delta} \in \mathbf{R}^{N\times N}$ and $\mathbf{E} \in \mathbb{C}^{N\times N}$, $\boldsymbol{\Delta}$ lower triangular Toeplitz, and $\|\boldsymbol{\Delta}\| < \delta, \|\mathbf{E}\| < \varepsilon$. Rewrite $\mathbf{T}+\mathbf{E}$ as,

$$\begin{aligned} \mathbf{T}+\mathbf{E} &= \mathbf{X}+\mathbf{Y}\boldsymbol{\Delta}+\mathbf{E} \\ &= \mathbf{X}+\begin{bmatrix}\frac{\delta}{\varepsilon}\mathbf{Y} & \mathbf{I}\end{bmatrix}\mathbf{D}\begin{bmatrix}\mathbf{I}\\ \mathbf{I}\end{bmatrix}, \end{aligned} \tag{12}$$

where,

$$\mathbf{D} = \begin{bmatrix}\frac{\varepsilon}{\delta}\boldsymbol{\Delta} & 0\\ 0 & \mathbf{E}\end{bmatrix}. \tag{13}$$

Now, $\|\mathbf{D}\| < \varepsilon$, so, by Theorem 1,

$$\left\|\begin{bmatrix}\mathbf{I}\\ \mathbf{I}\end{bmatrix}(z\mathbf{I}-\mathbf{X})^{-1}\begin{bmatrix}\frac{\delta}{\varepsilon}\mathbf{Y} & \mathbf{I}\end{bmatrix}\right\| > \varepsilon^{-1}.$$

Therefore, it is also true that $z \in \tilde{\sigma}_{\delta,\varepsilon}(\mathbf{T})$, from which it follows that $\sigma_{\delta,\varepsilon}(\mathbf{T}) \subset \tilde{\sigma}_{\delta,\varepsilon}(\mathbf{T})$. ■

The boundary of $\tilde{\sigma}_{\delta,\varepsilon}(\mathbf{T})$ can be computed in a similar manner to the pseudospectra (see the Appendix). However, the matrix $\mathbf{D}$ is structured in $\sigma_{\delta,\varepsilon}(\mathbf{T})$, but unstructured and complex for the set $\tilde{\sigma}_{\delta,\varepsilon}(\mathbf{T})$. Specifically, $\mathbf{D}$ in (13) is block diagonal and $\boldsymbol{\Delta}$ is real-valued and lower triangular. Thus, the boundaries of $\tilde{\sigma}_{\delta,\varepsilon}(\mathbf{T})$ are a conservative upper bound on the boundaries of $\sigma_{\delta,\varepsilon}(\mathbf{T})$. To obtain tight bounds, methods for incorporating structured perturbations in the modified pseudospectra computations need to be developed. An alternative lower bounding method is presented in the next section that does not rely on the spectral set bounding approach.

B. Approach 2: Random Uncertainty

The second approach to allowing uncertainty in transient growth analysis is to randomly select disturbances $\boldsymbol{\Delta}$ and compute the worst case pseudospectra of $\mathbf{T}$ for

the selected disturbances. In this way, a lower bounding set is computed for the modified pseudospectra $\sigma_{\delta,\varepsilon}(\mathbf{T})$.

The second approach can be summarized as follows.

1) Randomly select perturbations $\mathbf{\Delta}_l, l = 1,...,m$ with $\|\Delta_l\| = \delta$, as described below.
2) For each l, compute the pseudospectra of $\mathbf{T}_l = \mathbf{X} + \mathbf{Y}\mathbf{\Delta}_l$; specifically, compute the function $c_l(z) = \underline{\sigma}(z\mathbf{I} - \mathbf{T}_l) = 1/\left\|(z\mathbf{I} - \mathbf{T}_l)^{-1}\right\|$ at all points in a computational grid in a region of interest as described in the Appendix.
3) For the points in the computational grid, set $c(z) = \min_l c_l(z)$.
4) Plot the level sets of the function $\log_{10}\left(c(z)\right)$. A value of $-p$ for a level set of this function corresponds to a ε value of 10^{-p} for the worst case pseudospectra of $\mathbf{T}$ for the selected disturbances.

Remark: Disturbances with the largest possible norm, i.e., $\|\Delta\| = \delta$, are selected in this approach to best approximate the boundaries of the modified pseudospectra set $\sigma_{\delta,\varepsilon}(\mathbf{T})$ as closely as possible.

The random perturbations, $\mathbf{\Delta}_l$, are lower triangular Toeplitz, which can be written as

$$\mathbf{\Delta}_l = \begin{bmatrix} \xi_l(0) & 0 & \cdots & 0 \\ \xi_l(1) & \xi_l(0) & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ \xi_l(N-1) & \xi_l(N-2) & \cdots & \xi_l(0) \end{bmatrix},$$

for some $\xi_l(i)$, $i = 0,\ldots,N-1$. Therefore, one can select $\xi_l(i)$, $i = 0,\ldots,N-1$ randomly from any distribution and then normalize the set so that $\|\Delta_l\| = \delta$. Because $\mathbf{\Delta}_j$ represents a dynamic system perturbation [13],

$$\mathbf{\Delta}_l(z) = \xi_l(0) + \xi_l(1)z^{-1} + \cdots + \xi_l(N-1)z^{-(N-1)},$$

and band-limiting $\mathbf{\Delta}_l$ corresponds to truncating the perturbation impulse response, thereby emphasizing high-frequency perturbations over low-frequency perturbations.

Uncertainties for many physical systems tend to be largest at high frequencies, so high-frequency perturbations are expected to generate the worst case $\Delta_l(z)$.

Figure 4 shows the estimate for the modified pseudospectra $\sigma_{\delta,\varepsilon}(\mathbf{T})$ using the approach outlined above with $\delta = 1$. The modified pseudospectra is calculated for the system presented in Section V with weighting parameters q=10000, r=0, and s=1 (known to yield transient growth). Plot 4.a shows the modified pseudospectra calculated using 1000 random perturbations, which serves as a best estimate. The smallest level set to extend outside of the unit circle, $\varepsilon=10^{-4}$, is marked with an arrow. Plot 4.b shows that the pseudospectra can be well estimated using only a few perturbations by truncating Δ_l to emphasize high frequency perturbations. In particular, the smallest level set extending outside the unit circle is correctly obtained. Plots 4.c and 4.d demonstrate that a few random perturbations will not accurately estimate the pseudospectra when n is large, or high frequencies are not emphasized.

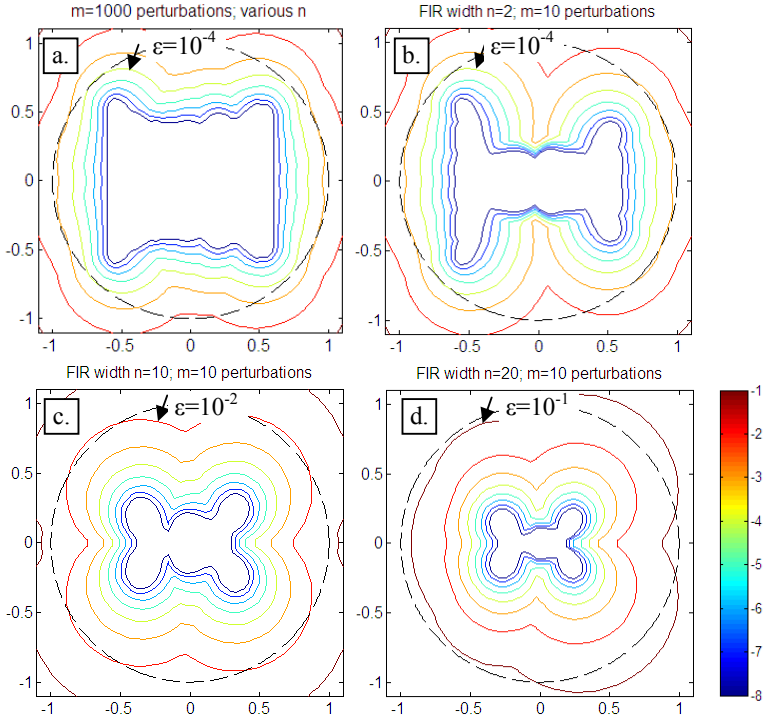

Figure 4. Modified pseudospectra for a perturbed set of systems, $\mathbf{T}_l = \mathbf{Q}(\mathbf{I} - \mathbf{L}\mathbf{P}_l)$, using Approach 2. Plot **a** shows the computed boundaries using 1000 random perturbations of $\mathbf{T}_l$ at several different Δ_l truncations. Plots **b, c, d** show the boundaries with 10 random perturbations and three different truncations. Arrows point to the largest ε level set that crosses the unit circle. The color bar is on a $\log_{10}$ scale so that the values of ε are $10^{-8},\ldots,10^{-1}$.

V. Applications in ILC Design

In this section a method is demonstrated by which pseudospectra may be used for ILC design. The goal here is to use pseudospectra analysis to predict whether a set of design parameters will yield small transient growth, and thus be acceptable for practical application. The approach presented here is based on a norm-optimal ILC design framework, but can be generalized for frequency domain [16] or LMI design of ILC. Of course, one needs to consider transient growth in norm-optimal ILC design only when the system is perturbed, so the results of Section IV.B are used to estimate the modified pseudospectra for a set of bounded system perturbations..

The simulations presented in the following two subsections use the following example system,

$$P(z) = C(zI - A)^{-1} B = \underbrace{P_{\text{model}}(z)}_{\substack{\text{low frequency} \\ \text{dynamics}}} \cdot \underbrace{\left(1 + W(z)\Delta(z)\right)}_{\substack{\text{uncertain high frequency} \\ \text{resonance}}},$$

with low frequency dynamics,

$$P_{\text{model}}(z) = \frac{0.0025(z+1)^2}{(z-1)(z-0.1)},$$

high frequency resonance filter,

$$W(z) = \frac{z^2 - 1.89z + 0.9}{z^2 - 1.6z + 0.8},$$

and unknown, bounded dynamics, $\|\Delta(z)\|_\infty < 1$. A Bode plot of $P(z)$ and the $P_{\text{model}}(z)$ is shown in Figure 5. A time horizon of N=60 is used.

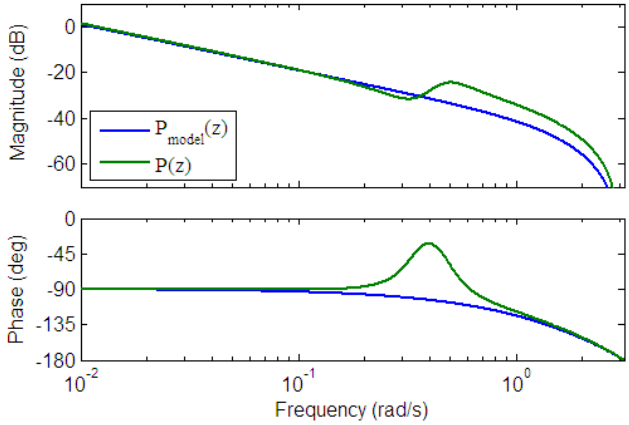

Figure 5. Bode plot of the nominal system $P_{model}(z)$ and perturbed system $P(z)$ for the Illustrative Example.

For the purposes of the control design, assume that $\Delta(z) = 0$ and use $P_{\text{model}}(z)$ as the nominal design model. The norm-optimal method [10] is used to design $\mathbf{Q}$ and $\mathbf{L}$ to minimize

$$J_{j+1} = \mathbf{e}_{j+1}^T \mathbf{Q} \mathbf{e}_{j+1} + \left(\mathbf{u}_{j+1} - \mathbf{u}_j\right)^T \mathbf{R}\left(\mathbf{u}_{j+1} - \mathbf{u}_j\right) + \mathbf{u}_{j+1}^T \mathbf{S} \mathbf{u}_{j+1}, \qquad (14)$$

for the design parameters positive definite $\mathbf{Q}$ and positive semi-definite $\mathbf{R}$ and $\mathbf{S}$. The resulting controller is well known and is given by

$$\begin{aligned}\mathbf{Q} &= \left(\mathbf{P}_\mathbf{m}^T\mathbf{Q}\mathbf{P}_\mathbf{m} + \mathbf{R} + \mathbf{S}\right)^{-1}\left(\mathbf{P}_\mathbf{m}^T\mathbf{Q}\mathbf{P}_\mathbf{m} + \mathbf{R}\right), \\ \mathbf{L} &= \left(\mathbf{P}_\mathbf{m}^T\mathbf{Q}\mathbf{P}_\mathbf{m} + \mathbf{R}\right)^{-1}\mathbf{P}_\mathbf{m}^T\mathbf{Q},\end{aligned} \tag{15}$$

where $\mathbf{P_m}$ is the lifted representation of $P_{\text{model}}(z)$. The state transition matrix is then given by

$$\mathbf{T} = \left(\mathbf{P}_\mathbf{m}^T\mathbf{Q}\mathbf{P}_\mathbf{m} + \mathbf{R} + \mathbf{S}\right)^{-1}\left(\mathbf{R} - \mathbf{P}_\mathbf{m}^T\mathbf{Q}\mathbf{P}_\mathbf{m}\mathbf{W}\boldsymbol{\Delta}\right),$$

where $\boldsymbol{\Delta}$ is the lifted representation of $\Delta(z)$ and $\mathbf{W}$ is the lifted representation of $W(z)$.

A. Learning Rate Selection

In this section, consider the simplified scalar weighting design, $\mathbf{Q}=q\mathbf{I}$, $\mathbf{R}=r\mathbf{I}$, $\mathbf{S}=s\mathbf{I}$. It is well known that the parameter r controls the transient response, but not the asymptotic solution. A small r results in fast convergence (for the nominal system), while a large r slows convergence. Although r is typically used to reduce noise sensitivity [17], here, its effect on transient growth is considered.

Consider the design parameters q=10000 and s=1, which result in transient growth when r=0. Figure 6 shows the modified pseudospectra calculated using Approach 2. The pseudospectra are generated using m=10 random perturbations with FIR support n=2. Increasing r has two predominant effects on the pseudospectra: 1) the eigenvalues (contained in the region encircled by the smallest level set) bunch together and approach one, and 2) the level sets shrink to a smaller radius around the eigenvalues, resulting in a smaller protrusion from the unit circle. The former corresponds to the well-known slowing of the learning as r increases, while the latter corresponds to a decrease in transient growth of the worst-case system in the perturbation set. The transient growth of one such perturbation, shown in Figure 7, illustrates these effects. By iterating over choices of r, an acceptable worst-case transient growth can be obtained.

One can study transient behavior more generally by examining the normality of the transition matrix. Rewriting $\mathbf{T}_l$ yields,

$$\mathbf{T}_l = \underbrace{\left(\mathbf{I} + \tfrac{q}{r}\mathbf{G}\right)^{-1}}_{\text{Normal}} + \underbrace{\left(\mathbf{I} + \tfrac{q}{r}\mathbf{G}\right)^{-1} \tfrac{q}{r}\mathbf{H}}_{\text{Non-normal}},$$

where $\mathbf{G} = \mathbf{P}_\mathbf{m}^T\mathbf{P}_\mathbf{m} + \frac{s}{q}\mathbf{I}$ and $\mathbf{H} = -\mathbf{P}_\mathbf{m}^T\mathbf{P}_\mathbf{m}\mathbf{W}\boldsymbol{\Delta}_l$. $\mathbf{T}_l$ is non-normal because $\mathbf{W}\boldsymbol{\Delta}_l$, and thus $\mathbf{H}$, are non-normal. However, as *r* approaches infinity, $\mathbf{T}_l$ approaches a normal matrix. Therefore, *r* acts as a robustifying variable in a practical sense by reducing transient growth for the worst-case system perturbations. The same is true as *q* approaches zero, but *q* affects asymptotic tracking performance, so it is not desirable to reduce.

Remark: The robustifying effect of *r* is only apparent in the case of transient growth. When *q* and *s* are selected to satisfy robust monotonic convergence conditions, *r* has little effect on robustness [14].

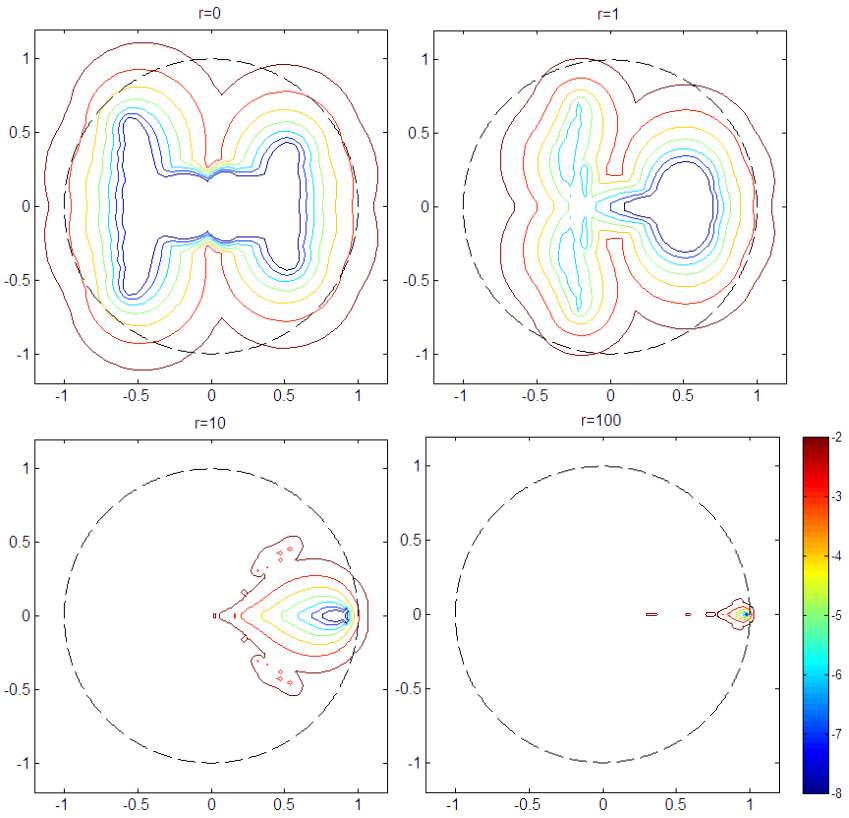

Figure 6. Modified pseudospectra for $\mathbf{T}_l$ with q=10000 and various values of r. The color bar is on a $\log_{10}$ scale so that the values of ε are $10^{-8},\ldots,10^{-1}$.

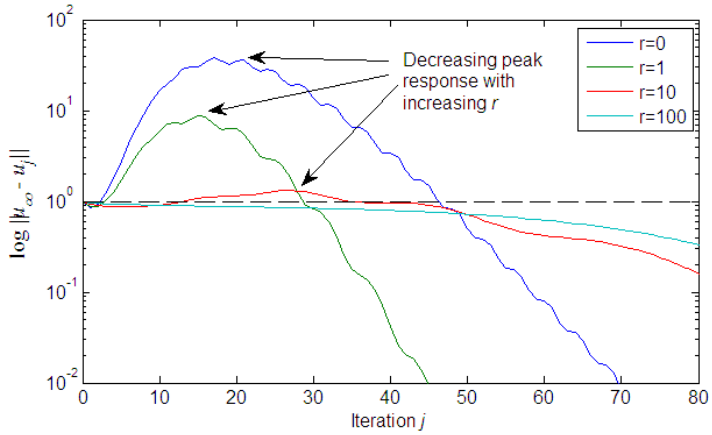

Figure 7. Transient growth for the $\mathbf{\Delta} = -\mathbf{I}$ perturbation with q=10000, s=1, and various values of r.

B. Time-Varying Q-filter Design

In this section consider the diagonal performance weighting, $\mathbf{Q} = diag\{q(0),\ldots,q(N-1)\}$. For simplicity, assume r=0 and s=1. The diagonal $\mathbf{Q}$ weighting is designed to emphasize performance over time segments, called α-

segments, where the error exhibits large time-frequency energy distributions [27], such as during step changes in the reference. This approach is similar to time-varying Q-filter bandwidth designs in the frequency domain [13,28], where it has been demonstrated that performance improvements are most significant when the learning system is not monotonic [29]. Here the method in [27] is extended for transient growth design.

The design procedure is as follows:

1. First, set $\mathbf{Q}=q^*\mathbf{I}$, where q^* is the largest value for which the transient growth is deemed acceptable. The value q^* can be obtained, for example, by tuning q on the physical system.
2. The pseudospectra is then calculated using Approach 2, and serves as the reference pseudospectra. Let ε^* denote the smallest ε-level set that extends outside the unit circle.
3. Replace $\mathbf{Q}=q^*\mathbf{I}$ with $\mathbf{Q} = diag\left\{q(0),\ldots,q(N-1)\right\}$ and iteratively shape $q(i)$, $i=0,...,N-1$ in the manner discussed in [27], with one difference. Rather than maintaining a particular robust monotonic bound as in [27], maintain the same ε^* as in the reference pseudospectra.

Consider, again, the example system. As a starting point, the scalar weighting $q=5000$ is treated as the largest weighting that yields acceptable transients. Learning transients and the pseudospectra are shown in Figure 8. Diagonal $\mathbf{Q}$ are then tuned by lowering q except during a segment around the time $k=20$ where the weighting is increased. Figure 8 shows two candidate designs for the diagonal $\mathbf{Q}$ along with their pseudospectra, transient growth profiles, and asymptotic tracking error. As can be seen, $\mathbf{Q}$ is reshaped without significantly altering the pseudospectra, and thus the transient response. In Design 2 a 20% reduction in weighting across most of the

iteration (from 5000 to 4000) allows a 40-fold increase of the weighting (from 5000 to 200000) during the critical "step", resulting in an overall 67% reduction in error.

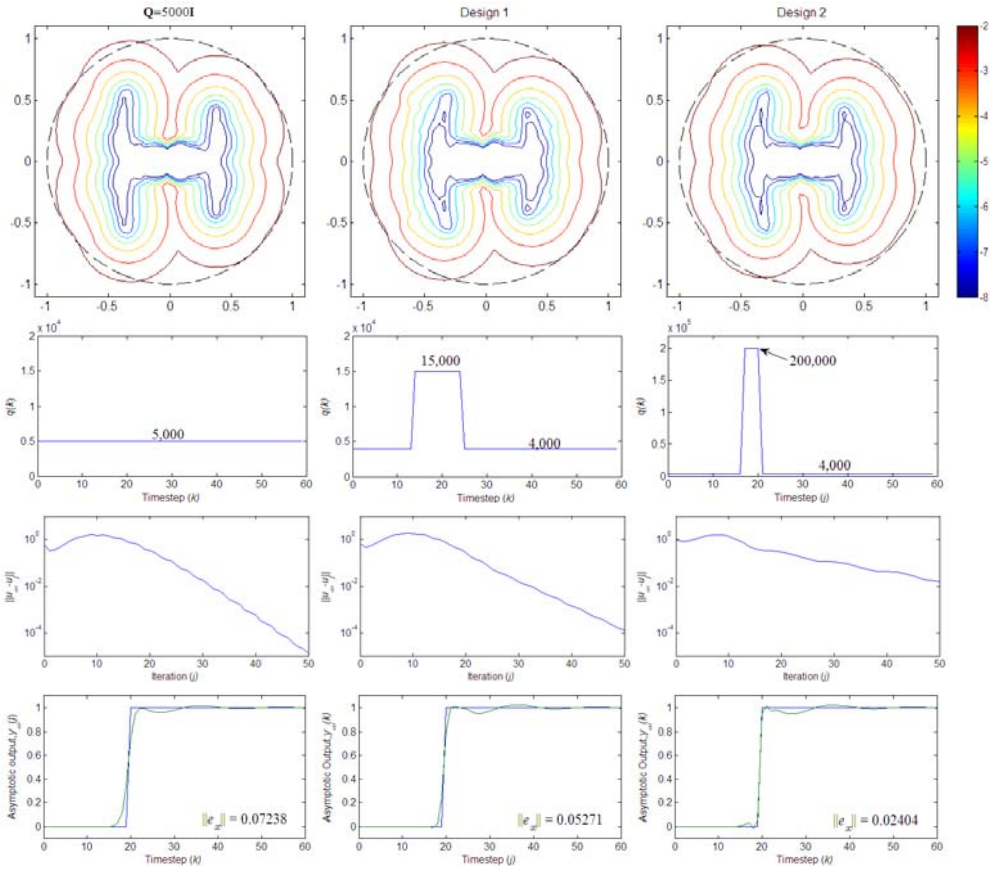

Figure 8. Pseudospectra, **Q** diagonal, transient growth, and asymptotic tracking for three designs of **Q**. The first column shows **Q**=5000**I**, column two shows Design 1 and column three shows Design 2.

VI. Discussion and Conclusions

This paper considers the problem of bounding transient growth in ILC systems. While transient growth is an undesirable property in ILC, robust monotonic convergence is a performance-limiting constraint. Thus, it is reasonable to consider some transient growth as a trade-off for improved performance when the system model is uncertain. The main contribution of this work is the extension of pseudospectra analysis to examine worst-case transient growth behavior in ILC systems with model uncertainty.

The proposed analysis is applied to two design problems in norm-optimal ILC: convergence rate parameter tuning and time-varying performance parameter design. In both problems, transient growth is limited while significantly improving asymptotic performance, despite model uncertainty. Thus, successful design in the transient growth regime is demonstrated.

Appendix: Computation of Pseudospectra

Computation of the pseudospectra is an actively researched topic [15]. Here, some main points are briefly addressed. First, note that Eigtool is a freely available Matlab program [30], which is widely used for pseudospectra computation. However, a modified pseudospectra computations is employed in this work. Therefore a basic pseudospectra algorithm is outlined as follows.

The preferred computational approach for pseudospectra is based on the characterization of the pseudospectra given above: a point z is in $\sigma_\varepsilon(\mathbf{T})$ if $\left\|(z\mathbf{I}-\mathbf{T})^{-1}\right\| > \varepsilon^{-1}$. Furthermore, it is known that z is in the boundary of the ε-pseudospectra if the above inequality is an equality. For the standard matrix 2-norm, $\left\|(z\mathbf{I}-\mathbf{T})^{-1}\right\| = 1/\underline{\sigma}(z\mathbf{I}-\mathbf{T})$, and therefore

$$\left\|(z\mathbf{I}-\mathbf{T})^{-1}\right\| = \varepsilon^{-1} \quad \Rightarrow \quad \underline{\sigma}(z\mathbf{I}-\mathbf{T}) = \varepsilon .$$

A simple algorithm for the computation of pseudospectra now proceeds as follows [15]:

1) Choose a region of interest in the complex plane and a computational grid for that region. (For the ILC pseudospectra problems in this work, the region of interest is a square surrounding the unit circle in the complex plane.)
2) At each point z in the computational grid, compute $c(z) = \underline{\sigma}(z\mathbf{I}-\mathbf{T})$.
3) Plot the level sets of the function $\log_{10}(c(z))$. A value of $-p$ for a level set of this function corresponds to a ε value of 10^{-p} for the matrix $\mathbf{T}$.

There are many ways to improve the computational efficiency of the basic pseudospectra algorithm. For example, computing the Schur decomposition of $\mathbf{T}$

gives $\mathbf{T} = \mathbf{USU}^*$, where **U** is unitary and **S** is an upper triangular matrix with the eigenvalues of **T** along the diagonal. Since the standard matrix 2-norm is unchanged by unitary transformations, $\left\|(z\mathbf{I}-\mathbf{T})^{-1}\right\| = \left\|(z\mathbf{I}-\mathbf{S})^{-1}\right\|$, and so

$$\left\|(z\mathbf{I}-\mathbf{T})^{-1}\right\| = \varepsilon^{-1} \quad \Rightarrow \quad \underline{\sigma}(z\mathbf{I}-\mathbf{S}) = \varepsilon .$$

Thus, the minimum singular value of an upper triangular matrix needs only be computed at each point in the computational grid. This minimum singular value computation can be performed efficiently using a Lanczos iteration; see [15] for details.

Remark: One problem that arises in lifted analysis of ILC systems is that the matrix size grows as $O(N^2)$ with the iteration length. For long iterations (greater than several thousand samples), computing $c(z) = \underline{\sigma}(z\mathbf{I}-\mathbf{T})$ on a regular PC is impractical due to memory limitations. One option is to use a Lanczos iteration [15] to calculate the singular value, but to replace every matrix calculation step with a dynamic simulation. This can be achieved by transforming **T** into the discrete-time system T(q) by reversing the lifting process. This approach was applied in [31] to calculate the maximum singular value of **T**, and the modification to calculate $c(z) = \underline{\sigma}(z\mathbf{I}-\mathbf{T})$ for the pseudospectra calculation is straightforward. Other methods for large-scale pseudospectra computations can be found in [15].

References

[1] Moore, K.L., Iterative Learning Control for Deterministic Systems, London: Springer-Verlag, 1993.

[2] Bristow, D.A., Tharayil, M., and Alleyne, A.G., "A Survey of Iterative Learning Control," *IEEE Control Systems Magazine*, vol. 26, no. 3, pp. 96-114, 2006.

[3] Ahn, H.-S., Chen, Y., and Moore, K.L., "Iterative Learning Control: Brief Survey and Categorization," *IEEE Transactions on Systems, Man, and Cybernetics-Part C*, vol. 37, no. 6, pp. 1099-121, 2007.

[4] Huang, Y.-C. and Longman, R.W., "Source of the Often Observed Property of Initial Convergence Followed by Divergence in Learning and Repetitive Control," *Advances in Astronautical Sciences*, vol. 90, no. 1, pp. 555-572, 1996.

[5] Longman, R.W., "Iterative Learning Control and Repetitive Control for Engineering Practice," *International Journal of Control*, vol. 73, no. 10, pp. 930-954, 2000.

[6] Longman, R.W. and Huang, Y.-C., "The Phenomenon of Apparent Convergence Followed by Divergence in Learning and Repetitive Control," *Intelligent Automation and Soft Computing*, vol. 8, no. 2, pp. 107-128, 2002.

[7] Chang, C.-K., Longman, R.W., and Phan, M., "Techniques for Improving Transients in Learning Control Systems," *Proceedings of the AAS/AIAA Astrodynamics Conference,* pp. 2035-2052, 1991 .

[8] Chen, H.-J. and Longman, R.W., "The Importance of Smooth Updates in Producing Good Error Levels in Repetitive Control," *Proceedings of IEEE Conference on Decision and Control,* pp. 258-63, 1999.

[9] Rogers, E., Galkowski, K., and Owens, D.H., Control Systems Theory and Applications for Linear Repetitive Processes Lecture Notes in Control and Information Sciences, Springer, 2007.

[10] Amann, N., Owens, D.H., and Rogers, E., "Iterative Learning Control for Discrete-Time Systems With Exponential Rate of Convergence," *IEE Proceedings: Control Theory and Applications*, vol. 143, no. 2, pp. 217-224, 1996.

[11] Harte, T.J., Hatonen, J., and Owens, D.H., "Discrete-Time Inverse Model-Based Iterative Learning Control: Stability, Monotonicity and Robustness," *International Journal of Control*, vol. 78, no. 8, pp. 577-586, 2005.

[12] Owens, D.H., Hatonen, J.J., and Daley, S., "Robust Monotone Gradient-Based Discrete-Time Iterative Learning Control," *International Journal of Robust and Nonlinear Control*, vol. 19, no. 6, pp. 634-661, 2009.

[13] Bristow, D.A. and Alleyne, A.G., "Monotonic Convergence of Iterative Learning Control for Uncertain Systems Using a Time-Varying Filter," *IEEE Transactions on Automatic Control*, vol. 53, no. 2, pp. 582-585, 2008.

[14] van de Wijdeven, J., Donkers, T., and Bosgra, O., "Iterative Learning Control for Uncertain Systems: Robust Monotonic Convergence Analysis," *Automatica*, vol. 45, no. 10, pp. 2383-2391, 2009.

[15] Trefethen, L.N. and Embree, M., Spectra and Pseudospectra, New Jersey: Princeton University Press, 2005.

[16] Bristow, D.A. and Singler, J.R., "Analysis of Transient Growth in Iterative Learning Control Using Pseudospectra," *Symposium on Learning Control,* Shanghai, China, 2009.

[17] Tousain, R., Van Der Meche, E., and Bosgra, O., "Design Strategy for Iterative Learning Control Based on Optimal Control," *Proceedings of the IEEE Conference on Decision and Control,* pp. 4463-4468, 2001.

[18] Hladowski, L., Cai, Z., Galkowski, K., Rogers, E., Freeman, C.T., Lewin, P.L., and Paszke, W., "Repetitive Process Based Iterative Learning Control Designed by LMIs and Experimentally Verified on a Gantry Robot," *Proceedings of the American Control Conference,* pp. 949-954, 2009.

[19] Amann, N., Owens, D.H., and Rogers, E., "Predictive Optimal Iterative Learning Control," *International Journal of Control*, vol. 69, no. 2, pp. 203-226, 1998.

[20] Trefethen, L.N. and Embree, M., Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators2005.

[21] Henrici, P., "Bounds for Iterates, Inverses, Spectral Variation and Fields of Values of Non-Normal Matrices," *Numerische Mathematik*, vol. 4, 1962.

[22] Gil', M.I., Norm Estimations for Operator-Valued Functions and Applications, New York: Marcel Dekker Inc., 1995.

[23] Golub, G.H. and Van Loan, C.F., Matrix Computations, Baltimore, MD: Johns Hopkins University Press, 1996.

[24] Hinrichsen, D. and Pritchard, A.J., Mathematical Systems Theory, Berlin: Springer-Verlag, 2005.

[25] Veselic, K., "Bounds for Exponentially Stable Semigroups," *Linear Algebra and Its Applications*, vol. 358, no. 1-3, pp. 309-333, 2003.

[26] Hinrichsen, D. and Kelb, B., "Spectral Value Sets: a Graphical Tool for Robustness Analysis," *Systems and Control Letters*, vol. 21, no. 2, pp. 127-136, 1993.

[27] Bristow, D.A., "Weighting Matrix Design for Robust Monotonic Convergence in Norm Optimal Iterative Learning Control," *Proceedings of the American Control Conference,* pp. 4554-60, 2008.

[28] Bristow, D.A., Dong, J., Alleyne, A.G., Ferreira, P., and Salapaka, S., "High Bandwidth Control of Precision Motion Instrumentation," *Review of Scientific Instruments*, vol. 79, no. 10, 2008.

[29] Bristow, D.A., Alleyne, A.G., and Tharayil, M., "Optimizing Learning Convergence Speed and Converged Error for Precision Motion Control," *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, vol. 130, no. 5, pp. 0545011-0545018, 2008.

[30] Wright, T.G., "EigTool," http://www.comlab.ox.ac.uk/pseudospectra/eigtool/, 2002.

[31] Barton, K.L., Bristow, D.A., and Alleyne, A.G., "A Numerical Method for Determining Monotonicity and Convergence Rate in Iterative Learning Control," *International Journal of Control*, vol. 83, no. 2, pp. 219-226, 2010.