Iterative learning control for spatio-temporal dynamics using Crank-Nicholson discretization

B. Cichy · K. Gałkowski · E. Rogers

© Springer Science+Business Media, LLC 2010

Abstract   Iterative learning control is now well established for linear and nonlinear dynamics in terms of both the underlying theory and experimental application. This approach is specifically targeted at cases where the same operation is repeated over a finite duration with resetting between successive repetitions. Each repetition or pass is known as a trial and the key idea is to use information from previous trials to update the control input used on the current one with the aim of improving performance from trial-to-trial. In this paper, new results on ILC applied to systems that arise from discretization of bi-variate partial differential equations describing spatio-temporal systems or processes are developed. Theses are based on Crank-Nicholson discretization of the governing partial differential equation, resulting in an unconditionally numerically stable approximation of the dynamics. It is also shown that this setting allows the selection of a finite number of points for sensing and actuation. The resulting control laws can be computed using Linear Matrix Inequalities (LMIs). Finally, an illustrative example is given and areas for further research are discussed.

Keywords   Iterative learning control · Crank-Nicholson discretization · Linear matrix inequalities

B. Cichy (✉) · K. Gałkowski
Institute of Control and Computation Engineering, University of Zielona Gora,
Podgorna 50, 65-246 Zielona Gora, Poland
e-mail: b.cichy@issi.uz.zgora.pl

K. Gałkowski
e-mail: k.galkowski@issi.uz.zgora.pl

E. Rogers
School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK
e-mail: etar@ecs.soton.ac.uk

Published online: 19 August 2010

 🕿 Springer

1 Introduction

Iterative Learning Control (ILC) has been especially developed to improve the performance of systems that operate in a repetitive manner where the task is to follow some specified trajectory in a specified finite time interval, also known as a pass or a trial in the literature, with high precision. The novel principle behind ILC is to suitably use information from previous trials, often in combination with appropriate current trial information, to select the current trial input to sequentially improve performance from trial-to-trial. In particular, the aim is to improve performance from trial-to-trial in the sense that the tracking error, the difference between the output on a trial and the specified reference trajectory, is sequentially reduced to either zero, in the ideal case, or some suitably small value. Since the original work Arimoto et al. (1984), the general area of ILC has been the subject of intense research effort. Initial sources for the literature here are the survey papers Bristow et al. (2006) and Ahn et al. (2007).

ILC algorithms propagate information from trial-to-trial and along the trial respectively and hence can be treated as a 2D system. In particular, there has been work on the use of a 2D discrete linear systems setting for the analysis and design of linear ILC control schemes based on the well known Roesser (1975) and Fornasini-Marchesini Fornasini & Marchesini (1978) state-space models, see, for example, Kurek & Zaremba (1993). These results focus entirely on trial-to-trial error convergence but it is known that a conflict can occur between the speed of the trial-to-trial error convergence and the performance produced along the trials. Since the trial length is finite then the response along any trial will be bounded, since even an unstable linear system can only produce a bounded response over such an interval and this could lead to unacceptable along the trial behavior, for example, exponential growth.

Repetitive processes are characterized Rogers et al. (2007) by a series of sweeps, termed passes, through a set of dynamics defined over a fixed finite duration known as the pass length. In particular, a pass is completed and then the process is reset before the start of the next one. On each pass, an output, termed the pass profile, is produced which acts as a forcing function on, and hence contributes to, the dynamics of the next pass profile. This, in turn, leads to the unique control problem where the output sequence of pass profiles generated can contain oscillations that increase in amplitude in the pass-to-pass direction. Industrial examples of these processes are detailed in Rogers et al. (2007). Recently, ILC algorithms designed in the repetitive process setting have been experimentally tested with results that clearly show how trial-to-trial error convergence and along the trial performance can be treated in this setting Hładowski et al. (2008), Hładowski et al. (2010).

Currently, the vast majority of the work reported on ILC considers finite-dimensional systems but there has been some work reported on its application to distributed parameter systems governed by Partial Differential Equations (PDEs), for example, Choi et al. (2001), Moore & Chen (2006), Qu (2002), Chao et al. (2009), Zhao (2005). In terms of developing ILC for PDEs, an obvious approach is to work directly with the defining equations, where, for example, Chao et al. (2009) considers the design of P-Type and D-Type control laws for parabolic PDEs, such as the controlled heat equation, using semigroup theory. See also Zhao (2005) where a number of other possible application areas are considered, such as velocity and tension control for axially moving materials and electrostatic microbridge actuators. Physical constraints in many distributed systems means that only boundary control can be used. However, distributed sensors/actuators have also a long history in numerous areas and more recent developments in supporting technologies have led to renewed activity into their effective application, see, for example, Maxwell & Asokanthan (2004), Zhao & Rahn (2007).

Natural connections exist between distributed parameter and multidimensional, or nD, systems and these motivate the analysis in this paper. In particular, the governing PDEs are first discretized by an implicit discretization scheme based on the Crank-Nicholson method Crank & Nicolson (1947) whose main advantage is that it is unconditionally numerically stable. Hence, unlike explicit methods, additional numerical stability analysis of the discrete approximation is not required. Also this approach allows actuation and sensing over finite regions, which is clearly of relevance for many applications.

Once the discretized model is constructed, it is written as a discrete linear repetitive process state-space model which is then used for the development of control laws that can be computed using Linear Matrix Inequalities (LMIs). A numerical example using the heat equation is given to illustrate the results obtained and we begin in the next section with a summary of the necessary background on repetitive processes. Throughout this paper $M \succ 0$ (respectively $\prec 0$) denotes a real symmetric positive (respectively negative) definite matrix. Also the null and identity matrices with compatible dimensions are denoted by 0 and I respectively.

2 Linear repetitive processes

The unique characteristic of a repetitive, or multipass Rogers et al. (2007), process is a series of sweeps, termed passes, through a set of dynamics defined over a fixed finite duration known as the pass length. In particular, a pass is completed and then the process is reset before the start of the next one. On each pass, an output, termed the pass profile, is produced which acts as a forcing function on, and hence contributes to, the dynamics of the next pass profile. This, in turn, leads to the unique control problem where the output sequence of pass profiles generated can contain oscillations that increase in amplitude in the pass-to-pass direction.

In order to explain how such a process arises in an industrial application, consider longwall coal cutting, see the relevant references cited in Rogers et al. (2007), where coal is extracted by hauling the cutting machine along the coal face riding on a semi-flexible conveyor. At the end of each pass, the machine is hauled back in reverse to the starting position and then the machine and conveyor is pushed forward to rest on the newly cut pass profile, that is, the height of the stone/coal interface about some fixed datum line. The control objective is to steer the cutting head such that the maximum amount of coal is extracted without penetrating the stone/coal interface at either the top or bottom of the coal seam. The basic geometry confirms that the previous pass profile critically influences the next one and hence long wall coal cutting is a repetitive process. The stability issue is that the undulations in the floor profiles build up from pass-to-pass and when excessive productive work must stop to enable them to be removed.

Consider the case of discrete dynamics along the pass and let $\alpha < \infty$ denote the pass length and $k \geq 0$ the pass number or index. Such processes evolve over the subset of the positive quadrant in the 2D plane defined by $\{(p, k) : 0 \leq p \leq \alpha - 1, k \geq 0\}$, and the most basic discrete linear repetitive process state-space model Rogers et al. (2007) has the following form

$$
\begin{aligned}
x_{k+1}(p + 1) &= A x_{k+1}(p) + B u_{k+1}(p) + B_0 y_k(p) \\
y_{k+1}(p) &= C x_{k+1}(p) + D u_{k+1}(p) + D_0 y_k(p)
\end{aligned}
\tag{1}
$$

Here on pass k, $x_k(p) \in \mathbb{R}^n$ is the state vector, $y_k(p) \in \mathbb{R}^m$ is the pass profile vector, and $u_k(p) \in \mathbb{R}^r$ is the vector of control inputs.

In order to complete the process description it is necessary to specify the boundary conditions, that is, the pass state initial vector sequence and the initial pass profile. The simplest form of these is

$$x_{k+1}(0) = d_{k+1}, \quad k \geq 0$$
$$y_0(p) = f(p), \quad 0 \leq p \leq \alpha - 1 \tag{2}$$

where the $n \times 1$ vector d_{k+1} has known constant entries and $f(p)$ is an $m \times 1$ vector whose entries are known functions of p.

Attempts to control these processes using standard (or 1D) systems theory and algorithms fail precisely because such an approach ignores their inherent 2D systems structure, that is, information propagation occurs from pass-to-pass, governed by the independent integer variable $k \geq 0$, and along a given pass, governed by the independent variable p, $0 \leq p \leq \alpha - 1$, and also the initial conditions are reset before the start of each new pass. To remove these deficiencies, a rigorous stability theory has been developed Rogers et al. (2007) based on an abstract model of the dynamics in a Banach space setting which includes a very large class of processes with linear dynamics and a constant pass length as special cases, including those described by (1) and (2). In terms of their dynamics, it is the pass-to-pass coupling, noting again their unique feature, which is critical. This can be written in the form $y_{k+1} = L_\alpha y_k$, where $y_k \in E_\alpha$, and L_α is a bounded linear operator mapping E_α into itself, and E_α is a Banach space with norm denoted by $|| \cdot ||$. In the case of examples described by (1) and (2), L_α is the convolution operator for a 1D discrete linear system with (state, input, output and direct feedthrough respectively) state-space model matrices $\{A, B_0, C, D_0\}$.

The stability theory for linear repetitive processes with constant pass length consists of two distinct concepts. Recalling the unique control problem, a natural definition of stability is to demand that a bounded initial pass profile y_0 produces a bounded sequence of pass profiles $\{y_k\}_{k \geq 1}$, where bounded is in terms of the norm on the underlying function space. Asymptotic stability, demands bounded-input bounded-output (BIBO) stability over the fixed finite pass length $\alpha > 0$. In terms of the abstract model, this requires the existence of finite real scalars $M_\alpha > 0$ and $\lambda_\alpha \in (0, 1)$ such that $||L_\alpha^k|| \leq M_\alpha \lambda_\alpha^k$, $k \geq 0$, where $|| \cdot ||$ denotes both the norm on the underlying function space and the induced operator norm.

For processes described by (1) and (2) it has been shown elsewhere, see, for example, Chapter 3 of Rogers et al. (2007), that this property holds if, and only if, all eigenvalues of the matrix D_0 have modulus strictly less than unity, written as $r(D_0) < 1$ where $r(\cdot)$ denotes the spectral radius of its matrix argument. Also asymptotic stability guarantees the existence of a limit profile denoted here by y_∞ and in physical terms means that after a sufficiently large number of passes the process converges in the pass-to-pass direction and produces the same pass profile. In the case of an asymptotically stable process described by (1) and (2), the limit profile is described by a 1D discrete linear systems state-space model with state matrix $A_{lp} = A + B_0(I - D_0)^{-1}C$.

Consider the special case of (1) and (2) when $A = -0.5$, $B = 0$, $B_0 = 0.5 + \beta$, $C = 1$, $D = 0$, $D_0 = 0$ and β is a real scalar such that $|\beta| \geq 1$. In this case the matrix $r(A_{lp}) = |\beta| \geq 1$ and the dynamics of the limit profile $y_\infty(p)$ are unstable along the pass. Asymptotic stability holds here because an unstable 1D discrete linear system can only produce a bounded response over the finite duration $0 \leq p \leq \alpha - 1$. To prevent this arising, the approach used is to demand the BIBO property for all possible values of the pass length α, including $\alpha = \infty$. In the abstract model this property requires the existence of finite real numbers $M_\infty > 0$ and $\lambda_\infty \in (0, 1)$, which are independent of the pass length α, such that $||L_\alpha^k|| \leq M_\infty \lambda_\infty^k$, $k \geq 0$. Numerous sets of conditions for this property are known such as the following.

Theorem 1 ([Rogers et al. (2007)](#)) *A discrete linear repetitive processes described by ([1](#)) and ([2](#)) is stable along the pass if, and only if, (i) $r(D_0) < 1$, (ii) $r(A) < 1$, and (iii) all eigenvalues of*

$$G(z) = C(zI - A)^{-1}B_0 + D_0$$

have modulus strictly less than unity for all $|z| = 1$.

It is condition (iii) in Theorem [1](#) that is violated for the simple example given above and, in particular, the intuitively obvious condition $r(A) < 1$ plus asymptotic stability is not strong enough. Despite the fact that the three stability conditions can be tested by direct application of standard linear systems tests, this result has not proved to be a basis for control law design to achieve stability along the pass and/or desired performance. An alternative approach that does lead to control law design algorithms is [Rogers et al. (2007)](#) to use a Lyapunov function approach accompanied by the use of the Linear Matrix Inequality (LMI) methods, and this is how the ILC designs are developed in this paper.

In the next section we introduce the Crank-Nicholson discretization for PDEs [Crank & Nicolson (1947)](#) by means of an example.

3 Crank-Nicholson discretization for PDEs

Consider the heat transfer equation

$$\frac{\partial x(t, w)}{\partial t} = \widehat{\alpha}^2 \frac{\partial^2 x(t, w)}{\partial w^2} + \delta u(t, w) \tag{3}$$

where $x(t, w)$ is the heat flow, $u(t, w)$ is an input variable, and t and w are the time and space variables respectively. Digitally based control law design and implementation requires the construction of an appropriate approximation of the dynamics by difference equations. If a direct discretization method is applied to spatio-temporal dynamics there is the need to ensure numerical stability by selection of the sampling period(s) used. An alternative is to use a unconditionally numerically stable discretization, where in this paper the Crank-Nicholson discretization method [Crank & Nicolson (1947)](#), [Rabenstein & Steffen (2009)](#) is used. Next we describe its application using the heat equation as an example.

Consider first the homogenous version of ([3](#)) obtained by setting $\delta = 0$ and introduce the approximations

$$x(t, w) \approx \frac{x_{l+1}(p) + x_l(p)}{2}$$

$$\frac{\partial x(t, w)}{\partial t} \approx \frac{x_{l+1}(p) - x_l(p)}{T}$$

$$\frac{\partial x(t, w)}{\partial w} \approx \frac{x_l(p + 1) - x_l(p - 1)}{h} + \frac{x_{l+1}(p + 1) - x_{l+1}(p - 1)}{h}$$

$$\frac{\partial^2 x(t, w)}{\partial w^2} \approx \frac{x_l(p + 1) - 2x_l(p) + x_l(p - 1)}{2h^2}$$
$$+ \frac{x_{l+1}(p + 1) - 2x_{l+1}(p) + x_{l+1}(p - 1)}{2h^2} \tag{4}$$

Then by routine manipulations the homogenous version of ([3](#)) over $l = 0, 1, \ldots, N$; $p = 0, 1, 2, \ldots, \alpha - 1$, is approximated by

$$A_1 x_{l+1}(p + 1) + B_1 x_{l+1}(p) + C_1 x_{l+1}(p - 1)$$
$$= A_2 x_l(p + 1) + B_2 x_l(p) + C_2 x_l(p - 1) \tag{5}$$

where

$$A_1 = -A_2 = C_1 = -C_2 = -\frac{\widehat{\alpha}^2}{2h^2}, \quad B_1 = \frac{1}{T} + \frac{\widehat{\alpha}^2}{h^2}, \quad B_2 = \frac{1}{T} - \frac{\widehat{\alpha}^2}{h^2}$$

with associated boundary conditions

$$\begin{aligned} x_0(p) &= g(p), \quad 0 \le p \le \alpha - 1 \\ x_l(-1) &= w_l, \ x_l(\alpha) = g_l, \quad l > 0 \end{aligned} \tag{6}$$

where $g(p)$ is an $n \times 1$ vector whose entries are known functions of p, g_l, w_l are $n \times 1$ vectors with known constant entries. Also the sequences $\{g_l\}$, $\{w_l\}$ are assumed to be bounded.

The Crank-Nicholson discretization is unconditionally numerically stable but is also necessary to consider its accuracy and, in particular, how many grid points should be considered to give an acceptable value of this quantity. This is a well studied problem in numerical analysis and clearly needs to be very carefully considered for each application. Here the heat equation is used as an illustration of the ILC algorithms developed and it is assumed that the discrete approximation used is sufficiently numerically accurate.

It is also necessary to consider the limiting case when the number of space and time grid points tends to infinity, where Lax-Richtmyer theory Strikwerda (2004) states that the discretization of a PDE converges to the true solution if the following two conditions hold:

- the discretization of each differential operator present is consistent, and
- the resulting difference equation is stable.

In terms of the first of these conditions, the discretizations of $\frac{\partial x(t,w)}{\partial t}$ and $\frac{\partial^2 x(t,w)}{\partial w^2}$ used are consistent, since when T and h are infinitely small the limits are first and second order derivatives. Also stability is guaranteed by Crank-Nicholson method which is unconditionally stable. Obviously, the exact density of the time and space grids much be considered for each application. Moreover, Lax-Richtmyer theory answers the question of whether or not the discretization converges to the true solution but does not give any information on the speed of convergence or what the error is for a given number of points.

The discrete approximation above is in the form of an implicit 2D equation that cannot be directly used to construct a discrete recursive model approximation to the process dynamics. Consequently introduce the stacking vectors

$$\mathcal{X}(l) = [\, x_l(0)^T, \ x_l(1)^T, \ \ldots, \ x_l(\alpha - 1)^T \,]^T \tag{7}$$

and for the boundary points

$$\mathcal{X}_B(l) = [\, x_l(-1)^T, \ x_l(\alpha)^T \,]^T \tag{8}$$

Then (5) can be rewritten as

$$\mathcal{A}_1 \mathcal{X}(l + 1) = \mathcal{A}_2 \mathcal{X}(l) - \mathcal{C}_1 \mathcal{X}_B(l + 1) + \mathcal{C}_2 \mathcal{X}_B(l) \tag{9}$$

where $x_l(-1) = w_l$, $x_l(\alpha) = g_l$ and $\mathcal{X}(0) = d$ are known boundary conditions and

$$\mathcal{A}_i = \begin{bmatrix} B_i & A_i & & & 0 \\ C_i & B_i & A_i & & \\ & \ddots & \ddots & \ddots & \\ & & C_i & B_i & A_i \\ 0 & & & C_i & B_i \end{bmatrix}, \ \mathcal{C}_i = \begin{bmatrix} C_i & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & A_i \end{bmatrix} \tag{10}$$

$i = 1, 2$. Note also that $\mathcal{A}_i$, $i = 1, 2$, are tri-diagonal block Toeplitz matrices.

In order to control the process dynamics, $\delta \neq 0$ must hold, and to achieve a causal control scheme we cannot apply to the input $u(t, w)$ the averaging procedure, as in (4) for $x(t, w)$, and instead it is approximated by $u(t, w) \approx u_l(p)$. The discrete approximation to the dynamics of (3) can now be written in the form

$$\mathcal{A}_1 \mathcal{X}(l+1) = \mathcal{A}_2 \mathcal{X}(l) - \mathcal{C}_1 \mathcal{X}_B(l+1) + \mathcal{C}_2 \mathcal{X}_B(l) + \mathcal{B}_1 \mathcal{U}(l) \tag{11}$$

where $\mathcal{U}(l) = [\, u_l(0)^T, \ u_l(1)^T, \ldots, u_l(\alpha - 1)^T \,]^T$, and $\mathcal{B}_1 = \delta I$. Finally, assuming the matrix inverse exists, left-multiply (11) by $\mathcal{A}_1^{-1}$ to obtain

$$\mathcal{X}(l+1) = \mathcal{A}\mathcal{X}(l) - \mathcal{D}_1 \mathcal{X}_B(l+1) + \mathcal{D}_2 \mathcal{X}_B(l) + \mathcal{B}\mathcal{U}(l) \tag{12}$$

where

$$\mathcal{A} = \mathcal{A}_1^{-1} \mathcal{A}_2, \quad \mathcal{B} = \mathcal{A}_1^{-1} \mathcal{B}_1, \quad \mathcal{D}_i = \mathcal{A}_1^{-1} \mathcal{C}_i, i = 1, 2 \tag{13}$$

and the boundary conditions are the same as for the homogenous case but left multiplied by the matrix $\mathcal{A}_1^{-1}$.

The construction of the state-space model (11) requires that the tri-diagonal Toeplitz matrix $\mathcal{A}_1$ is invertible, which always holds, see Yueh (2005). Even then, a very heavy computational load could result, especially for grids with a large spatial range, as inverse of the tri-diagonal matrix does not retain sparsity and also ill-conditioning may arise. Hence there is the need for further research that aims to extend the results given in this paper to apply directly to the descriptor model (11).

The Crank-Nicholson discretization scheme also allows the selection of spatial grid points where the signal values form an output vector

$$\mathcal{Y}(l) = \mathcal{C}\mathcal{X}(l) \tag{14}$$

and the entries in the matrix $\mathcal{C}$ are selected to reflect the grid points that contribute to the output vector. For example, if

$$\mathcal{Y}(l) = [\, x_l(0)^T, \ x_l(2)^T, \ \ldots, \ x_l(\alpha - 1)^T \,]^T \tag{15}$$

then

$$\mathcal{C} = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix} \tag{16}$$

It is important to note that in the finite-dimensional repetitive process state-space model the pass profile index is a spatial variable and the along the pass variable is temporal. In the discretization of a PDE these roles are reversed.

4 ILC Problem Formulation

In ILC for finite-dimensional systems, the successive executions of the task over the finite interval is termed a trial. Once a trial has been completed the process is reset to the same location prior to the start of the next one. During the resetting time the control input to be applied on the next trial is computed using information measured during the previous trial or trials and is often referred to as trial-to-trial updating. The overall objective is to improve performance from trial-to-trial by adjusting the input applied on each trial in such a way that,

at the very least, the error between the desired reference signal, say $r(p)$, and the output on any trial, say $y_k(p)$ where the integer $k \geq 0$ denotes the trial number, converges to zero or, more practically realistic, a suitably small value under an appropriate measure.

In general, two aspects of the performance of ILC control algorithms must be considered, namely $e_k(p) = r(p) - y_k(p)$ as k increases and $y_k(p)$ as p varies over the trial duration, and they can conflict since fast trial-to-trial error convergence may come at the cost of a trial output that has poor transient performance in the p direction, such as very lightly damped dynamics. The use of repetitive process theory to design ILC algorithms with experimental verification on a gantry robot is detailed in Hładowski et al. (2008), Hładowski et al. (2010) where this setting allows these two performance issues to be simultaneously considered.

In this section, the ILC problem considered is the development of control law design algorithms for discretized PDE model system where the reference is $y_l^*(p)$ for $l = 0, 1, \ldots N$, $p \in \mathcal{J}$, N is a positive integer, and $\mathcal{J} = \{j_0, j_1, \ldots, j_{\beta-1}\} \subset \{0, 1, \ldots, \alpha - 1\}$. The output vector is specified only on the sub-grid $\mathcal{J} = \{j_0, j_1, \ldots, j_{\beta-1}\}$ and at these points outputs signals are simply state signals, that is, for $l = 0, 1, \ldots N$; $p \in \mathcal{J}$ $y_k(p) = x_k(p)$ and the remaining grid points do not contribute to the output vector. If the state vectors at the selected grid points are denoted by $x_l^*(\cdot)$, they can be written in stacked vector form as $\mathcal{Y}^*(l) = [x_l^*(j_0)^T, x_l^*(j_1)^T, \ldots, x_l^*(j_{\beta-1})^T]^T$.

The spatial grid points at which actuation is applied are denoted by the sub-grid $\mathcal{I} = \{i_0, i_1, \ldots, i_{\gamma-1}\} \subset \{0, 1, \ldots, \alpha - 1\}$, that is, the input signal is defined over $\mathcal{I}$ as

$$\mathcal{U}(l) = [u_l(i_0)^T, u_l(i_1)^T, \ldots, u_l(i_{\gamma-1})^T]^T$$

Note that both the output, or sensing, and input, or actuating, sub-grids can be made equal to the full grid by selecting the index set as $\{0, 1, \ldots, \alpha - 1\}$, and in this case the output vector is equal to the state vector and the input vector is specified over the full grid.

Now assume that the N time steps define the dynamics along a trial. Once a trial is completed the process resets and the operation repeated, where the dynamics along the next trial is influenced by the previous trial values. Consider, for example, a bar of finite length where the temperature along its length is detected by α equi-spaced actuators and sensors. The design task is to produce a prescribed temperature profile by applying control over N time steps, then resetting before the next N time steps, recomputing the control signals making use of previous values, and so on. This sequence of operations has an ILC structure where the integer k denotes the trial number and the task then is to iterate in k to reduce error from trial-to-trial. Hence, the state-space model (12) and (14) can be written for the k-th trial as

$$\mathcal{X}(k, l+1) = \mathcal{A}\mathcal{X}(k, l) - \mathcal{D}_1 \mathcal{X}_B(k, l+1) + \mathcal{D}_2 \mathcal{X}_B(k, l) + \mathcal{B}\mathcal{U}(k, l)$$
$$\mathcal{Y}(k, l) = \mathcal{C}\mathcal{X}(k, l) \tag{17}$$

where the integer $k \geq 0$, $\mathcal{Y}(k, l) = [x_l(k, j_0)^T, x_l(k, j_1)^T, \ldots, x_l(k, j_{\beta-1})^T]^T$ and

$$\mathcal{X}(k, l) = [x_l(k, 0)^T, x_l(k, 1)^T, \ldots, x_l(k, \alpha - 1)^T]^T$$
$$\mathcal{U}(k, l) = [u_l(k, 0)^T, u_l(k, 1)^T, \ldots, u_l(k, \alpha - 1)^T]^T$$
$$\mathcal{X}_B(k, l) = [x_l(k, -1)^T, x_l(k, \alpha)^T]^T$$

Introduce

$$E(k, l) \cong \mathcal{Y}^*(l) - \mathcal{Y}(k, l), \quad 0 \leq l \leq N \tag{18}$$

and hence

$$E(k+1, l) - E(k, l) = -(\mathcal{Y}(k+1, l) - \mathcal{Y}(k, l)) \tag{19}$$

Then

$$\Upsilon(k+1, l+1) = \mathcal{X}(k+1, l) - \mathcal{X}(k, l)$$
$$\Delta U(k+1, l) = U(k+1, l) - U(k, l) \tag{20}$$
$$\Upsilon(k+1, l+1) = \mathcal{A}\Upsilon(k+1, l) + \mathcal{B}\Delta U(k+1, l-1) - \mathcal{D}_1\Theta_B(k+1, l)$$
$$+ \mathcal{D}_2\Theta_B(k+1, l-1) \tag{21}$$

where $\Theta_B(k+1, l) = \mathcal{X}_B(k+1, l) - \mathcal{X}_B(k, l)$.

Consider the application of a control law of the form

$$\Delta U(k+1, l) = K_1\Upsilon(k+1, l+1) + K_2 E(k, l+1) \tag{22}$$

to (21) and hence

$$\Upsilon(k+1, l+1) = (\mathcal{A} + \mathcal{B}K_1)\Upsilon(k+1, l) + \mathcal{B}K_2 E(k, l) - \mathcal{D}_1\Theta_B(k+1, l)$$
$$+ \mathcal{D}_2\Theta_B(k+1, l-1) \tag{23}$$

Also (19) and (20) give

$$E(k+1, l) - E(k, l) = -\mathcal{C}\mathcal{A}\Upsilon(k+1, l) - \mathcal{C}\mathcal{B}\Delta U(k+1, l-1)$$
$$+ \mathcal{C}\mathcal{D}_1\Theta_B(k+1, l) - \mathcal{C}\mathcal{D}_2\Theta_B(k+1, l-1)$$

or, using (22),

$$E(k+1, l) = -\mathcal{C}(\mathcal{A} + \mathcal{B}K_1)\Upsilon(k+1, l) + (I - \mathcal{C}\mathcal{B}K_2)E(k, l)$$
$$+ \mathcal{C}\mathcal{D}_1\Theta_B(k+1, l) - \mathcal{C}\mathcal{D}_2\Theta_B(k+1, l-1) \tag{24}$$

Finally, introduce

$$\hat{A} = \mathcal{A} + \mathcal{B}K_1, \quad \hat{B}_0 = \mathcal{B}K_2, \quad \hat{C} = -\mathcal{C}(\mathcal{A} + \mathcal{B}K_1), \quad \hat{D}_0 = I - \mathcal{C}\mathcal{B}K_2 \tag{25}$$

to write (23) and (24) in the form

$$\Upsilon(k+1, l+1) = \hat{A}\Upsilon(k+1, l) + \hat{B}_0 E(k, l) - \mathcal{D}_1\Theta_B(k+1, l) + \mathcal{D}_2\Theta_B(k+1, l-1)$$
$$E(k+1, l) = \hat{C}\Upsilon(k+1, l) + \hat{D}_0 E(k, l) + \mathcal{C}\mathcal{D}_1\Theta_B(k+1, l)$$
$$- \mathcal{C}\mathcal{D}_2\Theta_B(k+1, l-1) \tag{26}$$

The state-space model (26) has the form of the discrete linear repetitive process state-space model (1), except for the additional terms Θ_B arising from the boundary conditions. Suppose that $\mathcal{X}_B(k, l)$ is bounded and that the boundary values $x_l(-1) = w_l$, $x_l(\alpha) = g_l$, $l > 0$ of (6) can be obtained by extrapolation. Then the stability theory for linear repetitive processes can be used to design the control law in the same manner as in Hładowski et al. (2008), Hładowski et al. (2010) for a finite-dimensional discrete linear systems state-space model.

5 ILC Design

Applying the condition for asymptotic stability to (26) gives $r(\hat{D}_0) = r(I - CBK_2) < 1$. This condition is precisely that obtained by applying Roesser state-space model 2D discrete linear systems stability theory to (26) as first proposed in Kurek & Zaremba (1993) to ensure trial-to-trial error convergence only. By following the discussion in Sect. 2, it is easy to construct examples where $r(\hat{D}_0) < 1$ but the performance along the trial is very poor. For example, Hładowski et al. (2008), Hładowski et al. (2010) give such a design for a gantry

robot whose axes have been modeled by frequency response tests. To prevent such problems stability along the trial should be used where a computationally feasible alternative to the conditions of Theorem 1 is to use a Lyapunov function approach where the computations are LMI based.

5.1 The case when the output vector is a sub-vector of a full state vector

For an example described by (26), take the candidate Lyapunov function as

$$V(k,l) = V_1(k,l) + V_2(k,l)$$

with

$$V_1(k,l) = \Upsilon(k+1,l)^T P_1 \Upsilon(k+1,l), \quad V_2(k,l) = E(k,l)^T P_2 E(k,l)$$

where $P_i \succ 0$, $i = 1, 2$, with associated increment

$$\Delta V(k,l) = V_1(k,l+1) - V_1(k,l) + V_2(k+1,l) - V_2(k,l)$$

It now follows, by direct application of results in Chapter 9 of Rogers et al. (2007) for discrete linear repetitive processes, that stability along the trial holds if $\Delta V(k,l) < 0$ for all k and l, which is equivalent to the requirement that

$$\Phi^T P \Phi - P \prec 0 \tag{27}$$

where

$$\Phi = \begin{bmatrix} \hat{A} & \hat{B}_0 \\ \hat{C} & \hat{D}_0 \end{bmatrix}, \quad P = \text{diag}\{P_1, P_2\}$$

and (27) is the 2D Lyapunov matrix inequality.

It is possible to develop (27) into an LMI condition and obtain the following result for stability along the trial under control action together with formulas for computing the control law matrices.

Theorem 2 *An ILC scheme of the form (26) is stable along the trial over $R = \{(l, p) : l = 0, 1, \ldots, N; \ p = 0, 1, \ldots, \alpha - 1\}$ for any choice of the positive integers N and $\alpha > 1$ if there exist matrices $X_1 \succ 0$, $X_2 \succ 0$, R_1 and R_2 such that the following LMI is feasible*

$$M = \begin{bmatrix} -X_1 & 0 & X_1 \mathcal{A}^T + R_1^T \mathcal{B}^T & -X_1 \mathcal{A}^T \mathcal{C}^T - R_1^T \mathcal{B}^T \mathcal{C}^T \\ 0 & -X_2 & R_2^T \mathcal{B}^T & X_2 - R_2^T \mathcal{B}^T \mathcal{C}^T \\ \mathcal{A}X_1 + \mathcal{B}R_1 & \mathcal{B}R_2 & -X_1 & 0 \\ -\mathcal{C}\mathcal{A}X_1 - \mathcal{C}\mathcal{B}R_1 & X_2 - \mathcal{C}\mathcal{B}R_2 & 0 & -X_2 \end{bmatrix} \prec 0 \tag{28}$$

If (28) holds, control law matrices K_1 and K_2 can be computed using

$$K_1 = R_1 X_1^{-1}, \quad K_2 = R_2 X_2^{-1} \tag{29}$$

Proof By the 2D Lyapunov inequality, (26) is stable along the trial if there exists $P = \text{diag}\{P_1, P_2\} \succ 0$ such that

$$\Phi^T P \Phi - P \prec 0 \tag{30}$$

where

$$\Phi = \begin{bmatrix} \hat{A} & \hat{B}_0 \\ \hat{C} & \hat{D}_0 \end{bmatrix} \tag{31}$$

An obvious application of the Schur's complement formula to (30) yields

$$
\begin{bmatrix}
-P_1 & \star & \star & \star \\
0 & -P_2 & \star & \star \\
\hat{A} & \hat{B}_0 & -P_1^{-1} & \star \\
\hat{C} & \hat{D}_0 & 0 & -P_2^{-1}
\end{bmatrix} \prec 0
\tag{32}
$$

where $\star$ denotes a symmetric block entry in a matrix. Now introduce

$$
X_1 = P_1^{-1}, \quad X_2 = P_2^{-1}
\tag{33}
$$

and pre- and post-multiply (32) by diag$\{X_1, X_2, I, I\}$ to obtain

$$
\begin{bmatrix}
-X_1 & \star & \star & \star \\
0 & -X_2 & \star & \star \\
\hat{A}X_1 & \hat{B}_0 X_2 & -X_1 & \star \\
\hat{C}X_1 & \hat{D}_0 X_2 & 0 & -X_2
\end{bmatrix} \prec 0
\tag{34}
$$

Use of (25) gives, after some routine manipulations,

$$
\begin{bmatrix}
-X_1 & \star & \star & \star \\
0 & -X_2 & \star & \star \\
\mathcal{A}X_1 + \mathcal{B}K_1 X_1 & \mathcal{B}K_2 X_2 & -X_1 & \star \\
-\mathcal{C}\mathcal{A}X_1 - \mathcal{C}\mathcal{B}K_1 X_1 & X_2 - \mathcal{C}\mathcal{B}K_2 X_2 & 0 & -X_2
\end{bmatrix} \prec 0
\tag{35}
$$

Finally, let

$$
R_1 = K_1 X_1, \quad R_2 = K_2 X_2
\tag{36}
$$

to obtain (28), and the control law matrices which define (29) can be calculated from (36). This completes the proof. □

The following corollary of this last result, where a particular form is assumed for the decision matrices R_1, X_1, R_2, and X_2, increases computational efficiency.

Corollary 1 *An ILC scheme of the form (26) is stable along the trial over $R = \{(l, p) : l = 0, 1, \ldots, N; \ p = 0, 1, \ldots, \alpha - 1\}$ for any choice of the positive integers N and $\alpha > 1$ if there exist matrices $\hat{X}_i \succ 0$, $\hat{R}_{ij}, i = 1, 2, j = 1, 2, 3$, such that the following LMI is feasible*

$$
\begin{bmatrix}
-X & Y^T \\
Y & -X
\end{bmatrix} \prec 0
\tag{37}
$$

where

$$
X = \operatorname{diag}\{X_1, X_2\}, \quad X_1 = I_\alpha \otimes \hat{X}_1, \quad X_2 = I_{\lceil \alpha/2 \rceil} \otimes \hat{X}_2
$$
$$
Y = \begin{bmatrix}
\mathcal{A}X_1 + \mathcal{B}R_1 & \mathcal{B}R_2 \\
-\mathcal{C}\mathcal{A}X_1 - \mathcal{C}\mathcal{B}R_1 & X_2 - \mathcal{C}\mathcal{B}R_2
\end{bmatrix}
\tag{38}
$$

$\otimes$ *denotes the matrix Kronecker product, $\lceil \alpha/2 \rceil = \alpha/2$ if α is even and $\lceil \alpha/2 \rceil = \frac{\alpha+1}{2}$ if α is odd, and R_i, $i = 1, 2$, is defined as*

$$R_1 = \begin{bmatrix} \hat{R}_{12} & \hat{R}_{13} & & & 0 \\ \hat{R}_{11} & \hat{R}_{12} & \hat{R}_{13} & & \\ & \ddots & \ddots & \ddots & \\ & & \hat{R}_{11} & \hat{R}_{12} & \hat{R}_{13} \\ 0 & & & \hat{R}_{11} & \hat{R}_{12} \end{bmatrix}, \quad R_2 = \begin{bmatrix} \hat{R}_{22} & 0 & 0 & \dots & 0 & 0 \\ \hat{R}_{21} & \hat{R}_{23} & 0 & & & \\ 0 & \hat{R}_{22} & 0 & & & \vdots \\ & \hat{R}_{21} & \hat{R}_{23} & & & \vdots \\ & 0 & \hat{R}_{22} & \ddots & 0 & \\ \vdots & & \hat{R}_{21} & \ddots & \hat{R}_{23} & \\ \vdots & & 0 & \ddots & \hat{R}_{22} & 0 \\ \vdots & & & & \hat{R}_{21} & \hat{R}_{23} \\ 0 & 0 & 0 & \dots & 0 & \hat{R}_{22} \end{bmatrix} \tag{39}$$

If (37) holds, control law matrices K_1 and K_2 can be computed using

$$K_i = R_i X_i^{-1}, \quad i = 1, 2 \tag{40}$$

and hence are of the following form

$$K_1 = \begin{bmatrix} \hat{K}_{12} & \hat{K}_{13} & & & 0 \\ \hat{K}_{11} & \hat{K}_{12} & \hat{K}_{13} & & \\ & \ddots & \ddots & \ddots & \\ & & \hat{K}_{11} & \hat{K}_{12} & \hat{K}_{13} \\ 0 & & & \hat{K}_{11} & \hat{K}_{12} \end{bmatrix}, \quad K_2 = \begin{bmatrix} \hat{K}_{22} & 0 & 0 & \dots & 0 & 0 \\ \hat{K}_{21} & \hat{K}_{23} & 0 & & & \\ 0 & \hat{K}_{22} & 0 & & & \vdots \\ & \hat{K}_{21} & \hat{K}_{23} & & & \vdots \\ & 0 & \hat{K}_{22} & \ddots & 0 & \\ \vdots & & \hat{K}_{21} & \ddots & \hat{K}_{23} & \\ \vdots & & 0 & \ddots & \hat{K}_{22} & 0 \\ \vdots & & & & \hat{K}_{21} & \hat{K}_{23} \\ 0 & 0 & 0 & \dots & 0 & \hat{K}_{22} \end{bmatrix} \tag{41}$$

that is

$$\hat{K}_{ij} = \hat{R}_{ij} \hat{X}_i^{-1}, \quad i = 1, 2; \quad j = 1, 2, 3$$

The results in Theorem 2 and its corollary extend in a natural manner to other choices for the control and output matrices $\mathcal{B}$ and $\mathcal{C}$ respectively.

5.2 The case when the output vector is the full state vector

Consider now the special case when $\mathcal{C} = I$, that is, the output vector $\mathcal{Y}(k, l)$ is the full state vector $\mathcal{X}(k, l)$ and, for simplicity, zero boundary conditions. In this case, the result of Corollary 1 can be applied but this would introduce redundancy. Instead, consider the control law

$$\Delta U(k + 1, l) = \mathcal{K}_1 \Upsilon(k + 1, l) + \mathcal{K}_2 E(k, l + 1) \tag{42}$$

for which the ILC scheme (26) reduces to

$$E(k + 1, l + 1) = \widehat{A} \Upsilon(k + 1, l) + \widehat{F} E(k, l + 1) \tag{43}$$

where

$$\widehat{A} = -\mathcal{A} - \mathcal{B}\mathcal{K}_1, \quad \widehat{F} = I - \mathcal{B}\mathcal{K}_2 \tag{44}$$

Also

$$\Upsilon(k+1, l) = -(E(k+1, l) - E(k, l)) \tag{45}$$

and hence

$$E(k+1, l+1) = -\widehat{A}E(k+1, l) + \widehat{A}E(k, l) + \widehat{F}E(k, l+1)$$

which is a particular case of the Fornasini & Marchesini (1978) state-space model for 2D discrete linear systems. Hence the special case of $\mathcal{C} = I$ can equally well be analyzed using the theory for this well studied state-space model.

Using the Fornasini-Marchesini state-space model description, consider the candidate Lyapunov function

$$V'(k+1, l+1) = E(k+1, l+1)^T (\mathcal{P}_1 + \mathcal{P}_2) E(k+1, l+1) \tag{46}$$

where $\mathcal{P}_1 \succ 0$ and $\mathcal{P}_2 \succ 0$ and associated increment

$$\Delta V(k, l) = V'(k+1, l+1) - V''(k, l+1) - V_2(k+1, l) \tag{47}$$

where

$$V''(k, l+1) = E(k, l+1)^T \mathcal{P}_2 E(k, l+1) \tag{48}$$

and

$$V_2(k+1, l) := \Upsilon(k+1, l)^T \mathcal{P}_1 \Upsilon(k+1, l) \tag{49}$$

Lemma 1 *A 2D discrete linear system described by the Fornasini Marchesini state-space model* (43) *is stable over* $R = \{(l, p) : l = 0, 1, \ldots, N; \ p = 0, 1, \ldots, \alpha - 1\}$ *for any choice of the positive integers* N *and* $\alpha > 1$ *if there exists* $\mathcal{P}_1 \succ 0$ *and* $\mathcal{P}_2 \succ 0$ *such that*

$$\begin{bmatrix} \widehat{A}^T \mathcal{P}_1 \widehat{A} + \widehat{A}^T \mathcal{P}_2 \widehat{A} - \mathcal{P}_1 & \widehat{A}^T \mathcal{P}_1 \widehat{F} + \widehat{A}^T \mathcal{P}_2 \widehat{F} \\ \widehat{F}^T \mathcal{P}_1 \widehat{A} + \widehat{F}^T \mathcal{P}_2 \widehat{A} & \widehat{F}^T \mathcal{P}_1 \widehat{F} + \widehat{F}^T \mathcal{P}_2 \widehat{F} - \mathcal{P}_2 \end{bmatrix} \prec 0 \tag{50}$$

Proof The inequality $\Delta V(k, l) < 0$ can be rewritten as

$$\begin{bmatrix} \Upsilon(k+1, l) \\ E(k, l+1) \end{bmatrix}^T \left(\begin{bmatrix} \widehat{A} & \widehat{F} \\ \widehat{A} & \widehat{F} \end{bmatrix}^T \begin{bmatrix} \mathcal{P}_1 & 0 \\ 0 & \mathcal{P}_2 \end{bmatrix} \begin{bmatrix} \widehat{A} & \widehat{F} \\ \widehat{A} & \widehat{F} \end{bmatrix} - \begin{bmatrix} \mathcal{P}_1 & 0 \\ 0 & \mathcal{P}_2 \end{bmatrix} \right) \begin{bmatrix} \Upsilon(k+1, l) \\ E(k, l+1) \end{bmatrix} < 0$$

which holds for all non-zero $\begin{bmatrix} \Upsilon(k+1, l) \\ E(k, l+1) \end{bmatrix}$ when

$$\begin{bmatrix} \widehat{A} & \widehat{F} \\ \widehat{A} & \widehat{F} \end{bmatrix}^T \begin{bmatrix} \mathcal{P}_1 & 0 \\ 0 & \mathcal{P}_2 \end{bmatrix} \begin{bmatrix} \widehat{A} & \widehat{F} \\ \widehat{A} & \widehat{F} \end{bmatrix} - \begin{bmatrix} \mathcal{P}_1 & 0 \\ 0 & \mathcal{P}_2 \end{bmatrix} \prec 0 \tag{51}$$

This is equivalent to (50) and the proof is complete. $\qquad \square$

Lemma 1 does not provide a computationally feasible method for control law design as its stability condition is not in LMI form. Instead, it provides a basis for the following result.

Theorem 3 *A 2D discrete linear system described by the Fornasini-Marchesini state-space model (43) is stable over $R = \{(l, p) : l = 0, 1, \ldots, N;\ p = 0, 1, \ldots, \alpha - 1\}$ for any choice of the positive integers N and $\alpha > 1$ if there exists matrices $N_{ij},\ j = 1, 2, 3,\ \mathcal{P}_i \succ 0,\ i = 1, 2$, such that*

$$\begin{bmatrix} -\mathcal{P} & Y^T \\ Y & -\mathcal{P} \end{bmatrix} \prec 0 \tag{52}$$

where

$$\mathcal{P} = \mathrm{diag}\{\mathcal{P}_1, \mathcal{P}_2\}, \quad \mathcal{P}_i = I_\alpha \otimes P_i \tag{53}$$

$$Y = \begin{bmatrix} -\mathcal{A}\mathcal{P}_1 - \mathcal{B}\mathcal{N}_1 & \mathcal{P}_2 - \mathcal{B}\mathcal{N}_2 \\ -\mathcal{A}\mathcal{P}_1 - \mathcal{B}\mathcal{N}_1 & \mathcal{P}_2 - \mathcal{B}\mathcal{N}_2 \end{bmatrix} \tag{54}$$

$$\mathcal{N}_i = \begin{bmatrix} N_{i2} & N_{i3} & & & 0 \\ N_{i1} & N_{i2} & N_{i3} & & \\ & \ddots & \ddots & \ddots & \\ & & N_{i1} & N_{i2} & N_{i3} \\ 0 & & & N_{i1} & N_{i2} \end{bmatrix} \tag{55}$$

If (52) holds, stabilizing control law matrices in (42) are given by

$$\mathcal{K}_i = \mathcal{N}_i \mathcal{P}_i^{-1}, \quad i = 1, 2 \tag{56}$$

$$\mathcal{K}_i = \begin{bmatrix} K_{i2} & K_{i3} & & & 0 \\ K_{i1} & K_{i2} & K_{i3} & & \\ & \ddots & \ddots & \ddots & \\ & & K_{i1} & K_{i2} & K_{i3} \\ 0 & & & K_{i1} & K_{i2} \end{bmatrix}, \quad i = 1, 2 \tag{57}$$

Proof First, introduce

$$M = \begin{bmatrix} \widehat{A} & \widehat{F} \\ \widehat{A} & \widehat{F} \end{bmatrix} \tag{58}$$

and apply the Schur's complement formula to (51) to obtain, after pre- and post-multiplication by $P^{-1} \oplus I$,

$$\begin{bmatrix} -P^{-1} & P^{-1}M^T \\ MP^{-1} & -P^{-1} \end{bmatrix} \prec 0 \tag{59}$$

Setting

$$\mathcal{P} = P^{-1} \tag{60}$$

gives

$$\begin{bmatrix} -\mathcal{P} & Y^T \\ Y & -\mathcal{P} \end{bmatrix} \prec 0 \tag{61}$$

where

$$Y = M\mathcal{P} \tag{62}$$

Now expand the product (62) as

$$Y = \begin{bmatrix} \widehat{A} & \widehat{F} \\ \widehat{A} & \widehat{F} \end{bmatrix} \begin{bmatrix} \mathcal{P}_1 & 0 \\ 0 & \mathcal{P}_2 \end{bmatrix} = \begin{bmatrix} \widehat{A}\mathcal{P}_1 & \widehat{F}\mathcal{P}_2 \\ \widehat{A}\mathcal{P}_1 & \widehat{F}\mathcal{P}_2 \end{bmatrix} \tag{63}$$

and substitute for $\widehat{A}$ and $\widehat{F}$ from (44) to obtain

$$Y = \begin{bmatrix} -\mathcal{A}\mathcal{P}_1 - \mathcal{B}\mathcal{K}_1\mathcal{P}_1 & \mathcal{P}_2 - \mathcal{B}\mathcal{K}_2\mathcal{P}_2 \\ -\mathcal{A}\mathcal{P}_1 - \mathcal{B}\mathcal{K}_1\mathcal{P}_1 & \mathcal{P}_2 - \mathcal{B}\mathcal{K}_2\mathcal{P}_2 \end{bmatrix} \tag{64}$$

Setting

$$\mathcal{N}_1 = \mathcal{K}_1\mathcal{P}_1, \quad \mathcal{N}_2 = \mathcal{K}_2\mathcal{P}_2 \tag{65}$$

gives (54) and the proof is complete. □

This result is simpler than that of Corollary 1 but only in the case when $\mathcal{C} = I$.

A critical feature in the use of Crank-Nicholson discretization method in this paper is that lifting must be used to deal with the spatial variable. This, in turn, means the ILC scheme can be generalized to the case where not all space points are the subject of control action or where the control action is piecewise constant or sparse. These cases are of direct practical relevance and are detailed next.

5.2.1 Spatially piecewise constant and sparse controllers

In the general case when the control input is defined for all space points of the grid, the input matrix $\mathcal{B}_1$ in the discretized model satisfies $\mathcal{B}_1 = \delta I$. In at least some applications, control action will only be possible to be piecewise constant. For example, when the control signal has the same value at every three consecutive points $\{1, 2, 3\}, \{4, 5, 6\}, \ldots \ldots, \{77, 78, 79\}$ of the space grid, the matrix $\mathcal{B}_1$ has the following form

$$\mathcal{B}_1 = \begin{bmatrix} 0 & \delta & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \delta & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \delta & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta & 0 \\ & & & & \vdots & & & & \end{bmatrix} \cdots \tag{66}$$

It is also possible that control action can only be applied at selected space points of the grid. For example, when control is applied at points $1, 4, 7, 10, 13, \ldots, 79$ of the space grid,

the matrix $\mathcal{B}_1$ has the following form

$$\mathcal{B}_1 = \begin{bmatrix} 0 & \delta & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & \vdots \end{bmatrix} \tag{67}$$

In such a case, the control law (22) leads to overly conservative LMI conditions. An alternative control law is

$$\Delta U(k+1,l) = \mathcal{K}_1 E(k+1,l) + \mathcal{K}_2 E(k,l) + \mathcal{K}_3 E(k,l+1) \tag{68}$$

with corresponding controlled process state-space model

$$E(k+1,l+1) = \widehat{A}_1 E(k+1,l) - \widehat{A}_2 E(k,l) + \widehat{F} E(k,l+1) \tag{69}$$

where

$$\widehat{A}_1 = \mathcal{A} - \mathcal{B}\mathcal{K}_1, \quad \widehat{A}_2 = \mathcal{A} + \mathcal{B}\mathcal{K}_2, \quad \widehat{F} = I - \mathcal{B}\mathcal{K}_3 \tag{70}$$

This approximation of the spatio-temporal dynamics again has the structure of a 2D discrete linear system described by a Fornasini-Marchesini state-space model. Assume, without loss of generality, that the boundary conditions are zero and change the Lyapunov function (first introduced for (26)) increment to

$$\Delta V(k,l) = V(k+1,l+1) - V(k+1,l) - V(k,l) - V(k,l+1) \tag{71}$$

Then we have the following result.

Theorem 4 *A 2D discrete linear system described by (69) is stable over $R = \{(l,p) : l = 0, 1, \ldots, N; \ p = 0, 1, \ldots, \alpha - 1\}$ for any choice of the positive integers N and $\alpha > 1$ if there exists matrices $Q_1 \succ 0, \mathcal{N}_i, i = 1, 2, 3,$ of the form*

$$\mathcal{N}_i = \begin{bmatrix} N_{i2} & N_{i3} & & & 0 \\ N_{i1} & N_{i2} & N_{i3} & & \\ & \ddots & \ddots & \ddots & \\ & & N_{i1} & N_{i2} & N_{i3} \\ 0 & & & N_{i1} & N_{i2} \end{bmatrix} \tag{72}$$

such that the following LMI is feasible

$$\begin{bmatrix} -\widehat{Q} & Y^T \\ Y & -Q \end{bmatrix} \prec 0 \tag{73}$$

where

$$Q = I \otimes Q_1, \ \widehat{Q} = \mathrm{diag}\{Q, Q, Q\}, \ Y = \begin{bmatrix} \mathcal{A}Q - \mathcal{B}\mathcal{N}_1 & -\mathcal{A}Q - \mathcal{B}\mathcal{N}_2 & Q - \mathcal{B}\mathcal{N}_3 \end{bmatrix}$$

If ([73](#)) holds, stabilizing control law matrices are given by

$$\mathcal{K}_i = \mathcal{N}_i Q^{-1}, \quad i = 1, 2, 3 \tag{74}$$

where

$$\mathcal{K}_i = \begin{bmatrix} K_{i2} & K_{i3} & & & 0 \\ K_{i1} & K_{i2} & K_{i3} & & \\ & \ddots & \ddots & \ddots & \\ & & K_{i1}; & K_{i2} & K_{i3} \\ 0 & & & K_{i1} & K_{i2} \end{bmatrix}, \quad i = 1, 2 \tag{75}$$

Proof The proof follows that of the previous case with routine changes to account for the structure of the matrix $\mathcal{B}_1$ in this case. Hence the details are omitted. □

Obviously, the spatially piece-wise constant and sparse control can also be applied to the case when $\mathcal{Y}$ is a sub-vector of $\mathcal{X}$ and is a form of output control.

6 Numerical examples

Consider the heat flow equation of ([3](#)) over $0 \le t \le 1950$ [sec] and $0 \le w \le 243$ [m] when $\widehat{\alpha} = 0.5, \delta = 5$ and apply the Crank-Nicholson discretization scheme with $T = 26 \,[sec]$, and $h = 3 \,[m]$, where these are feasible since the Crank-Nicholson discretization is unconditionally numerically stable. Hence $\alpha = 81$, $N = 75$ and the matrices in the model ([5](#)) are

$$A_1 = -A_2 = C_1 = -C_2 = -0.0139, \quad B_1 = 0.0662, \quad B_2 = 0.0107$$

6.1 The output as a selection of state vectors

Consider the case when the output vector is formed by selecting every other state vector along the spatial grid, that is, $\mathcal{J} = \{j_0, j_1, \ldots, j_{\beta-1}\} = \{0, 2, \ldots, \alpha - 1\}$. The LMIs of Corollary [1](#) are feasible in this case and solving them gives the control law matrices K_i

$$\hat{K}_{11} = \hat{K}_{13} - 2.7778 \times 10^{-3}, \quad \hat{K}_{12} = -2.1368 \times 10^{-3}$$
$$\hat{K}_{21} = \hat{K}_{23} = -1.6418 \times 10^{-3}, \quad \hat{K}_{22} = 7.8303 \times 10^{-3}$$

The desired, or reference, $\mathcal{Y}^*(k)$ is shown in Fig.[1](#), and only has nonzero values for the even natural numbers $0, 2, \ldots, 80$, and so does the error. If zero boundary conditions are assumed, the initial error for trial $l = 0$ is the same as the reference and Fig. [2](#) shows the error after 30 and 50 (left and right-hand plots respectively) trials, demonstrating that fast error convergence is possible with, see Fig. [3](#), acceptable levels of control action.

6.2 The piecewise constant and sparse cases

Consider the case when the full state is taken as the system output vector, that is $\mathcal{C} = I$, but control is defined only at the space points $1, 4, 7, 10, 13, \ldots, 79$ on the grid and hence the control input matrix is of the form ([66](#)). The reference signal is now defined on the full grid as shown in Fig. [4](#).

Fig. 1 Reference signal $\mathcal{Y}^*(l)$
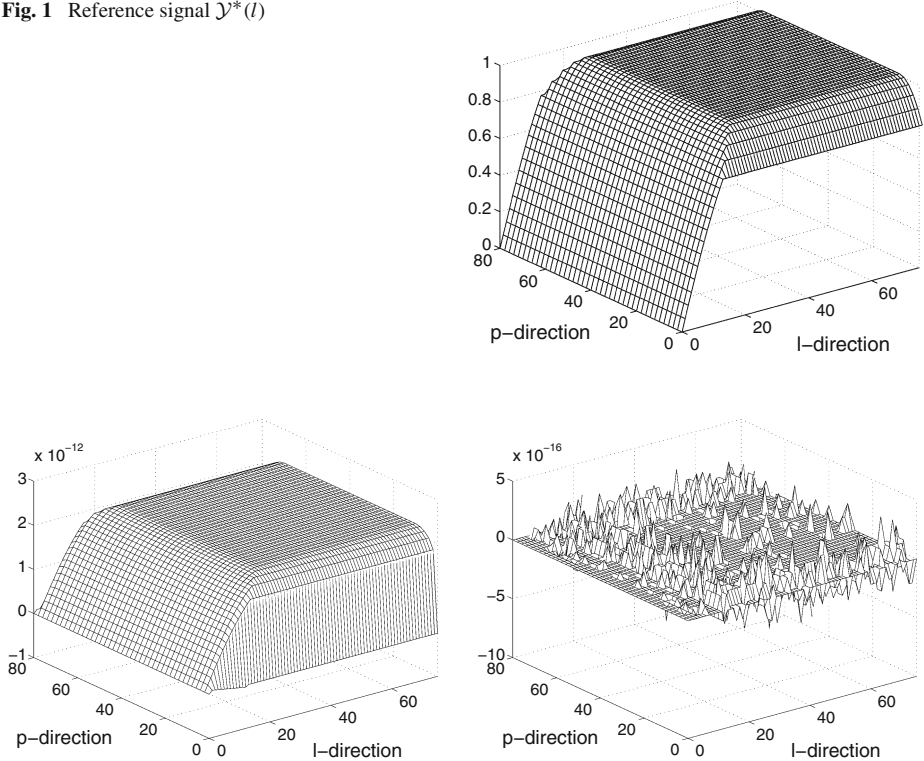
Fig. 2 Error dynamics after 30 (*left-hand plot*) and 50 (*right-hand plot*) trials
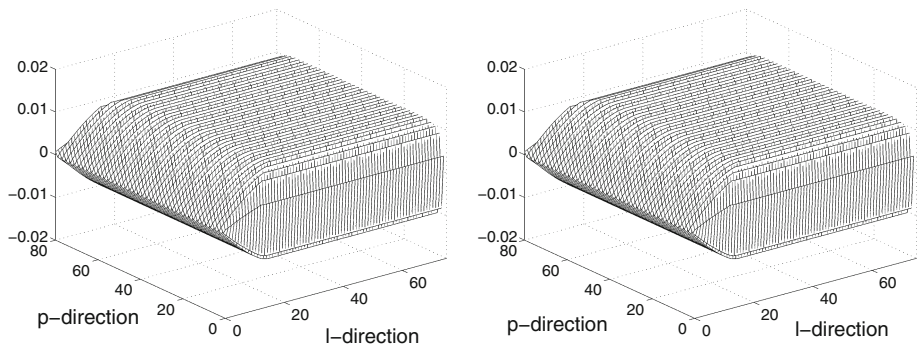
Fig. 3 Control input $\mathcal{U}(k, l)$ on trials 30 (*left-hand plot*) and 50 (*right-hand plot*)

The LMIs of Theorem 4 are feasible in this case and solving them gives

$$K_{11} = K_{13} = 5.263 \times 10^{-3}, \quad K_{12} = 2.6984 \times 10^{-3}$$
$$K_{21} = K_{23} = -5.263 \times 10^{-3}, \quad K_{22} = -2.6984 \times 10^{-3}$$
$$K_{31} = K_{33} = 1.4768 \times 10^{-3}, \quad K_{32} = -7.9343 \times 10^{-4}$$

Fig. 4 Reference signal $\mathcal{Y}^*(l)$
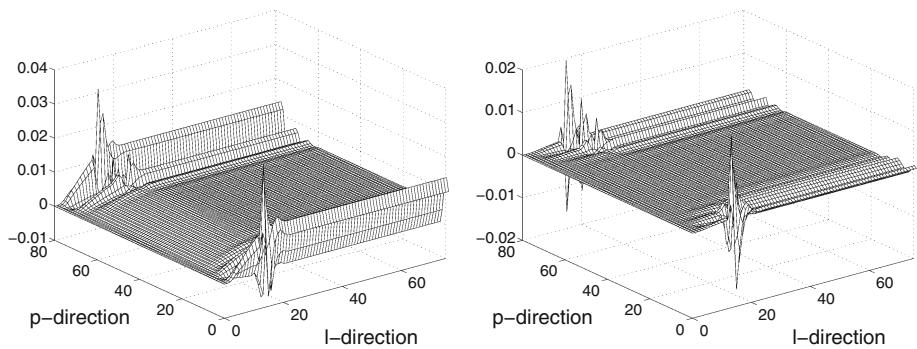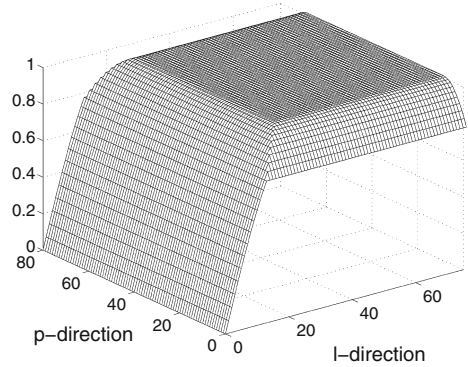
Fig. 5 Error dynamics on trials 30 (*left-hand plot*) and 50 (*right-hand plot*)
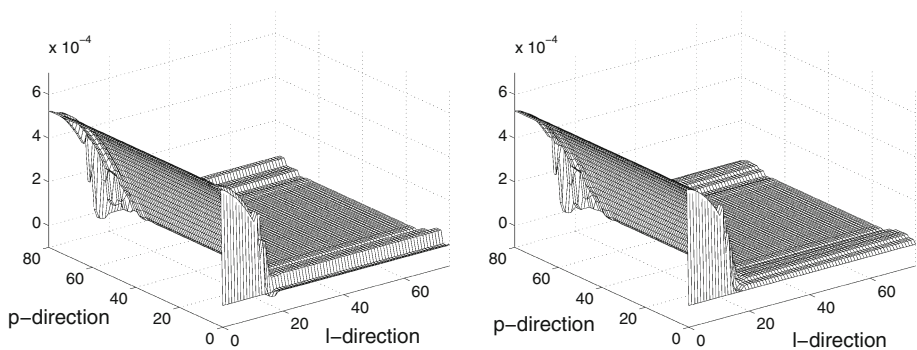
Fig. 6 Control input $\mathcal{U}(k, l)$ on trials 30 (*left-hand plot*) and 50 (*right-hand plot*)

which define the matrices (75) in the control law (68). Again zero boundary conditions are assumed, and hence the initial error is the same as the reference of Fig. 4. Figure 5 show the errors for trials $k = 30$ and $k = 50$ (left and right-hand plots respectively) and Fig. 6 the control action, which are acceptable.

6.2.1 The sparse case

Assume that the output is defined as in the previous numerical example, but the control is defined on, and applied at, the space grid points grid 1, 4, 7, 10, 13, ..., 79, resulting in a control matrix of the form (67). The reference is the same as in the previous case.

The LMIs of Theorem 4 are feasible and their solution yields the following matrices $K_{ij}, i = 1, 2, 3, j = 1, 2, 3$

$$K_{11} = -K_{21} = 0.0122, \quad K_{31} = 3.1511 \times 10^{-3}$$
$$K_{12} = -K_{22} = 3.4278 \times 10^{-3}, \quad K_{32} = 1.4572 \times 10^{-4}$$
$$K_{13} = -K_{23} = 7.2699 \times 10^{-4}, \quad K_{33} = 1.2445 \times 10^{-4}$$

which define the matrices (75) in the control law (68). Assume again zero boundary conditions, and hence the initial error (for trial $k = 0$) is the same as the reference of Fig. 4. Figure 7 shows the error for trials 30 (left-hand plot) and 50 (right-hand plot) respectively and Fig. 8 shows the control action. It is seen that the ILC scheme converges from trial-to-trial, which is further highlighted by the mean squared error data given in Fig. 9. Also the required input signal is acceptable.

In applications there will often be uncertainty associated with the model used for design and hence a robust control theory is an obvious area for further research. To give a preliminary
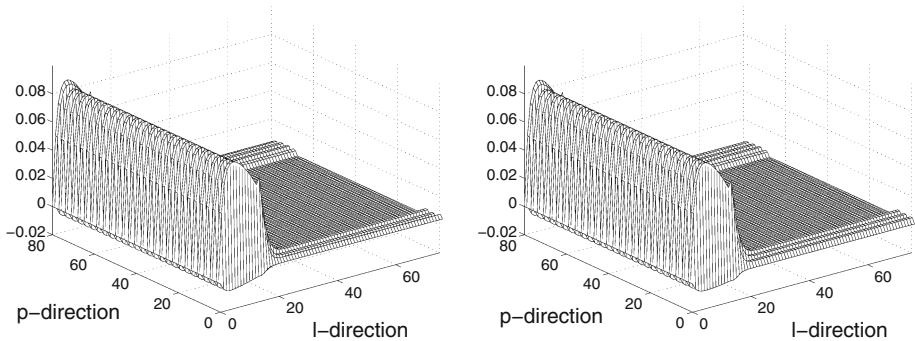

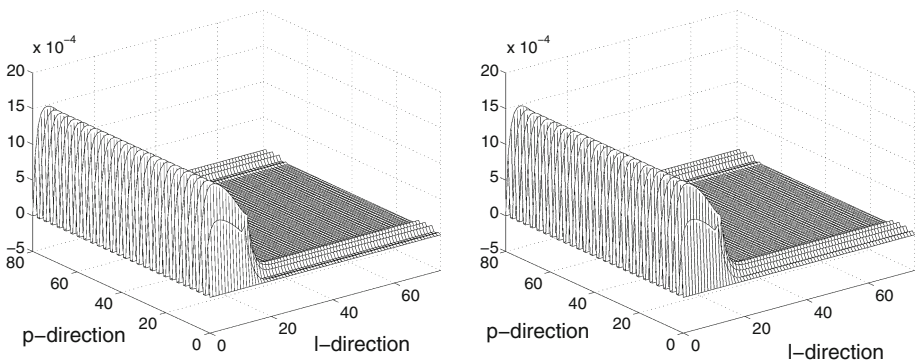**Fig. 7** Error dynamics on trials 30 (*left-hand plot*) and 50 (*right-hand plot*)

Fig. 8 Control input $\mathcal{U}(k, l)$ on trials 30 (*left-hand plot*) and 50 (*right-hand plot*)

Fig. 9 Mean square error
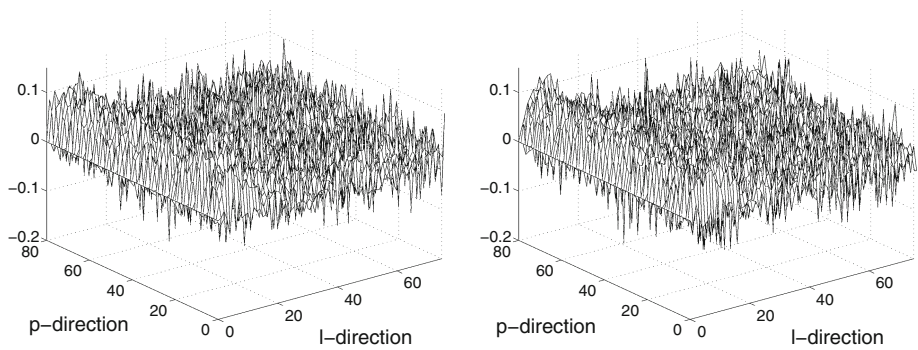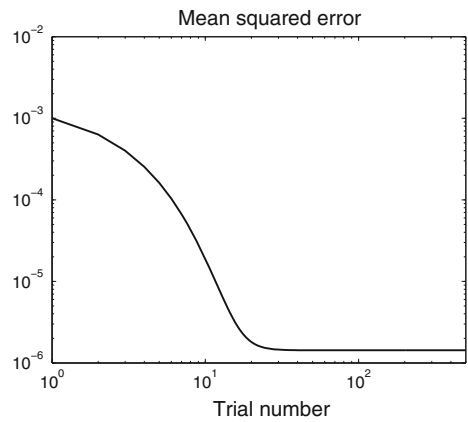dynamics in the trial-to-trial
direction

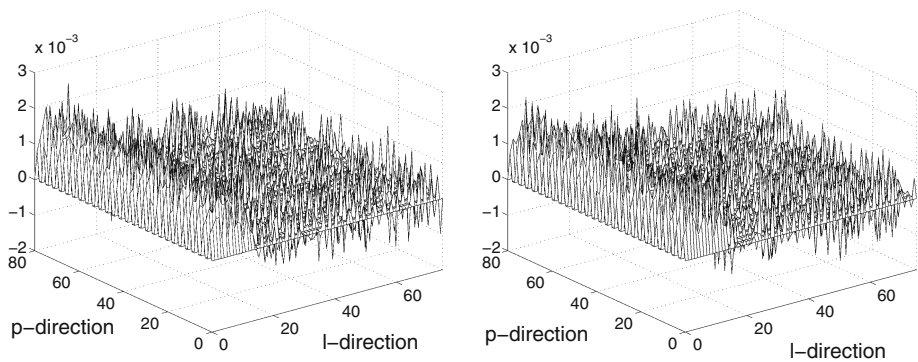**Fig. 10** Error dynamics after 30 (*left-hand plot*) and 50 (*right-hand plot*) trials

Fig. 11 Control input $\mathcal{U}(k, l)$ on trials 30 (*left-hand plot*) and 50 (*right-hand plot*)

assessment of control law performance in the presence of uncertainty, suppose that the control
law is designed on the data given above but simulated against the case when

$$A_1 = -A_2 = C_1 = -C_2 = -0.0135, \quad B_1 = 0.0643, \quad B_2 = 0.0104$$

Fig. 12 Mean square error dynamics in the trial-to-trial direction
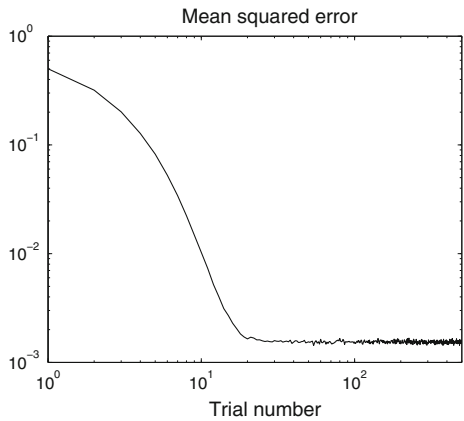
Also the system output is assumed to be corrupted by the subject to disturbances, modeled as additive white noise $N(0; 0.01)$ (normal distribution with mean 0 and standard deviation 0.01). Figures 10 and 12 show that, even though the overall error value increases slightly, the algorithm developed for a nominal system with no noise presence, works quite well. Also the control input values does not increase excessively, see Fig. 11.

7 Conclusions

In this paper we have developed new results on the application of the ILC to spatio-temporal systems described by partial differential equations using repetitive process and 2D linear systems settings for analysis. The heat transfer equation has been used as an example but the results are more general and can be easy extended to other PDEs with more space and time indeterminates that arise in applications, such as those detailed in Zhao (2005), Fiory (2001), Schaper et al. (1994)

The approach first requires discretization of the defining equations followed by the use of Lyapunov functions that satisfy sufficient but not necessary conditions for error convergence but do enable control law design to be undertaken using LMIs. This last feature offsets to some degree of the conservativeness arising from not using necessary and sufficient conditions.

The discretization method used in this work is the unconditionally stable Crank-Nicholson scheme. This leads to discrete linear repetitive process 2D discrete linear systems state-space models for design where the indeterminates are time and trial number respectively, and the space variable is 'hidden' in the internal process structure. The method used here also allows the control action to be applied to a subset of the spatio-temporal grid, which is of practical relevance. The extension to allow for boundary control only is the subject of ongoing work. Also this method allows the practically relevant case when the control input action is constant, or piecewise constant in the space domain, but variable in time.

In this work, we assume that a tri-diagonal block Toeplitz matrix in the Crank-Nicholson discretization approximation is to be inverted to obtain the state-space model required for control law design. This is a possible source of numerical problems as, although for the case considered in the paper the respective descriptor matrix ia always nonsingular, it is an obvious possible source of ill-conditioning. One way of avoiding this step is to use a descriptor model for design and this topic is currently under investigation.

References

Ahn, H.-S., Chen, Y. Q., & Moore, K. L. (2007). Iterative learning control: brief survey and categorization. *IEEE Transactions on Systems Man and Cybernetics Part C: Applications and Reviews, 37*(6), 1099–1121.

Arimoto, S., Kawamura, S., & Miyazaki, F. (1984). Bettering operation of robots by learning. *Journal of Robotic Systems, 1*(2), 123–140.

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

Chao, X., Arastoo, R., & Schuster, E. (2009). On iterative learning control of parabolic distributed parameter systems. In *Proceedings of the 17th mediterranean conference on control and automation* (pp. 510–515).

Choi, J., Seo, B. J., & Lee, K. S. (2001). Constrained digital regulation of hyperbolic PDE systems: A learning control approach. *Korean Journal of Chemical Engineering, 18*(5), 606–611.

Crank, J., & Nicolson, P. (1947). A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. *Proceedings of the Cambridge Philosophical Society, 43*, 50–67.

Fiory, A. T. (2001). Methods in microelectronics for rapid thermal annealing of implanted dopants. In B. L. Sopori (Ed.), 11th *Workshop on crystalline silicon solar cell materials and processes* (pp. 102–113).

Fornasini, E., & Marchesini, G. (1978). Doubly indexed dynamical systems: State-space models and structural properities. *Theory of Computing Systems, 12*(1), 59–72.

Hładowski, Ł., Cai, Z., Gałkowski, K., Rogers, E., Freeman, C.T., & Lewin, P. L. (2008). Using 2D systems theory to design output signal based iterative learning control laws with experimental verification. In *Proceedings of the 47th IEEE conference on decision and control* (pp. 3026–3031).

Hładowski, Ł., Gałkowski, K., Cai, Z., Rogers, E., Freeman, C. T., & Lewin, P. L. (2010). Experimentally supported 2D systems based iterative learning control law design for error convergence and performance. *Control Engineering Practice, 18*(4), 339–348.

Kurek, J. E., & Zaremba, M. B. (1993). Iterative learning control synthesis based on 2-D system theory. *IEEE Transactions on Automatic Control, 38*(1), 121–125.

Maxwell, N. D., & Asokanthan, S. F. (2004). Modal characteristics of a flexible beam with multiple distributed actuators. *Journal of Sound and Vibration, 269*(1–2), 19–31.

Moore, K. L., & Chen, Y. Q. (2006) Iterative learning approach to a diffusion control problem in an irrigation application. In *Proceedings of the 2006 IEEE international conference on mechatronics and automation* (pp. 1329–1334).

Qu, Z. (2002). An iterative learning algorithm for boundary control of a stretched moving string. *Automatica, 38*(5), 821–827.

Rabenstein, R., & Steffen, P. (2009). Implicit discretization of linear partial differential equations and repetitive processes. In *The 6th international workshop on multidimensional (nD) systems (nDS 2009)* (pp. 1–7).

Roesser, R. P. (1975). A discrete state-space model for linear image processing. *IEEE Transactions on Automatic Control, 20*(1), 1–10.

Rogers, E., Gałkowski, K., & Owens, D. H. (2007). Control systems theory and applications for linear repetitive processes, volume 349 of *Lecture notes in control and information sciences*. Springer.

Schaper, C. D., Moslehi, M. M., Saraswat, K. C., & Kailath, T. (1994). Modeling, identification, and control of rapid thermal processing systems. *Journal of the Electrochemical Society, 141*(11), 3200–3209.

Strikwerda, R. C. (2004). Finite difference schemes and partial differential equations, SIAM.

Yueh, W. C. (2005). Eigenvalues of several tridiagonal matrices. *Applied Mathematics E-Notes, 5*, 66–74.

Zhao, H. (2005). *Passive, iterative, and repetitive control for flexible distributed parameter systems*. PhD thesis, Department of Mechanical and Nuclear Engineering, The Pennsylvania State University.

Zhao, H., & Rahn, C. D. (2007). Stability of damped membranes and plates with distributed inputs. *Journal of Sound and Vibration, 302*(3), 564–576.

Author Biography

B. Cichy received the M.Sc. degree from the University of Zielona Gora, Poland in 2003. Currently, he is completing his Ph.D. research at the same university under the supervision of Prof. Krzysztof Galkowski. His research interests are on multidimensional (nD) systems, linear repetitive processes, iterative learning control, and the application of numerical methods to control problems. For his activity, he has been awarded the President of University of Zielona Gora Teamwork Second Class Award.