Iterative Learning Control based on Stochastic Approximation *

Mark Butcher, Alireza Karimi and Roland Longchamp

Automatic Control Laboratory, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland.
(e-mail: `alireza.karimi@epfl.ch`*)*

Abstract: In this paper stochastic approximation theory is used to produce Iterative Learning Control (ILC) algorithms which are less sensitive to stochastic disturbances, a typical problem for the learning process of standard ILC algorithms. Two algorithms are developed, one to obtain zero mean controlled error and one to minimise the mean 2-norm of the controlled error. The former requires a certain knowledge of the system but in the presence of noise can give reasonably rapid convergence. The latter can either use a model or be model free by employing a second experiment.

1. INTRODUCTION

Iterative Learning Control (ILC) is a technique used to enhance the tracking performance of systems that perform repetitive operations. Drawing on the principles of human learning, information 'learnt' from previous repetitions is used to improve the performance of the system during the next repetition/iteration i.e. reduce the tracking error. ILC has been shown to be very effective for systems that are predominately affected by deterministic, repetitive disturbances, which are learnt from one iteration to the next. However, when the system is affected by stochastic, non-repeating disturbances the tracking performance is greatly diminished (Gunnarsson and Norrlöf, 2006; Butcher et al., 2008). It is, therefore, important to develop ILC algorithms that have reduced sensitivity to their presence.

Although the deterministic aspects of ILC have received greater attention, certain researchers have already proposed algorithms which are robust to the presence of stochastic disturbances.

(1) The use of a forgetting factor in ILC was first proposed in Heinzinger et al. (1989) for a D-type ILC law. It was then proposed in Arimoto (1990) for P-type ILC. It is shown there that by introducing the forgetting factor the system's output converges to a neighbourhood of the desired one, despite the presence of norm-bounded initialisation errors, fluctuations of the dynamics and random disturbances. However, in Saab (2005) and Butcher et al. (2008), it is shown that, statistically, the use of a forgetting factor can increase the expected value and variance of the error signal, compared to standard ILC algorithms.

(2) The filtering of the ILC command has been proposed in certain papers as a way of reducing the influence of noise on the error (Norrlöf and Gunnarsson, 2001). However, whilst it reduces the error variance, it causes a nonzero converged expected value of the error.

(3) Kalman filtering-type techniques have also been applied to ILC to estimate the controlled output, in the presence of disturbances (Tao et al., 1994; Saab, 2001a,b; Norrlöf, 2002; Ahn et al., 2006). In the case of perfect knowledge of the disturbance covariance matrices and, for most methods, perfect system knowledge, convergence to optimal tracking, in a certain sense, can be shown. However, exact knowledge of these parameters is, typically, unrealistic.

(4) In Tao et al. (1994) another ILC algorithm is proposed using a learning gain that decreases each iteration and has the form of a Stochastic Approximation (SA) algorithm. No detailed analysis is, however, carried out. The application of SA theory to ILC is most directly considered in Chen (2003) and Chen and Fang (2004) for the linear and nonlinear cases respectively. It is shown that the proposed ILC law converges almost surely to the optimal input and the output error is minimised in the mean square sense as the number of iterations tend to infinity. The algorithm requires only that the optimal input is realizable. Neither knowledge of the disturbance covariance matrices nor the system is required. The disadvantage is slow convergence.

The aim of this paper is also to apply SA theory to ILC for linear systems affected by stochastic disturbances. Two methods are proposed. In the first, an algorithm is proposed to solve a root-finding type criterion to set the expected value of the controlled error equal to zero. In the second method, an ILC command to minimise a mean square criterion is developed. Conditions for almost sure convergence to the solution input are given for both methods. In addition, monotonic convergence of the error signal from one iteration to the next and robustness to system uncertainty are considered.

The proposed methods differ from that in Chen (2003) as there a simultaneous perturbation SA type algorithm is employed, which uses random perturbations to estimate the gradient. Here, either the use of a, possibly uncertain, system model or a second model-free experiment, using re-
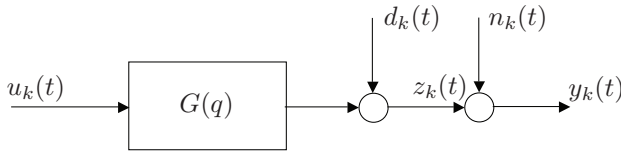
Fig. 1. System with a repetitive desired output

versed time inputs, is proposed. These proposed techniques will typically lead to faster convergence.

Steepest descent algorithms have been applied to ILC for the discrete-time case by Hatonen et al. (2003). Although certain similarities exist between the proposed SA based algorithm and steepest descent algorithms, the major difference is the conditions on the step sizes between iterations. These conditions are necessary for SA to assure almost sure convergence to the optimal input in the presence of stochastic disturbances. It is shown in this paper that the step sizes proposed in Hatonen et al. (2003) do not satisfy the SA conditions. Hence, whilst their algorithm provides rapid convergence to the optimal input in the deterministic case, this cannot be shown when stochastic disturbances are present.

This paper is organised as follows. In Section 2 the notational framework is defined and the assumptions are stated. In Section 3 Zero Mean Error ILC is proposed and analysed. Then in Section 4 the alternative criterion of minimising the mean 2-norm of the controlled error is examined. Finally in Section 5 conclusions are made.

2. PRELIMINARIES

2.1 Notation

We consider the SISO, linear, discrete-time, stable system $G(q)$, shown in Fig. 1, that carries out a repetitive task and whose controlled output, $z_k(t)$, at repetition k is given by:

$$z_k(t) = G(q)u_k(t) + d_k(t), \qquad (1)$$

where $u_k(t)$ is the input to the system, $d_k(t)$ the load disturbance and q the forward-shift, time domain operator. The system's measured output, $y_k(t)$, is:

$$y_k(t) = z_k(t) + n_k(t), \qquad (2)$$

where $n_k(t)$ is the measurement disturbance. It should be mentionned that if $G(q)$ represents a closed-loop transfer function then $d_k(t)$ and $n_k(t)$ will be the signals resulting from the filtering of external disturbances by the corresponding closed-loop transfer functions.

The controlled tracking error signal is defined as:

$$\epsilon_k(t) = y_d(t) - z_k(t), \qquad (3)$$

where $y_d(t)$ is the desired system output, which is defined over the finite repetition duration for $t = 0, \ldots, N-1$, and the measured error signal is given by:

$$e_k(t) = y_d(t) - y_k(t). \qquad (4)$$

As the signals are defined over a finite duration, it is possible to express the system's input-ouput relationship by a matrix representation. Taking advantage of the non-causal filtering possibilities of ILC, the lifted-system representation is typically used. For a system with a relative degree of m we define the vectors:

$$u_k = [u_k(0), u_k(1), \ldots, u_k(N-m-1)]^T \text{ and} \qquad (5)$$

$$z_k = [z_k(m), z_k(m+1), \ldots, z_k(N-1)]^T, \qquad (6)$$

with y_k, d_k, n_k and y_d defined similarly to z_k. Using these vectors, the measured output of the system is:

$$y_k = Gu_k + d_k + n_k, \qquad (7)$$

where G is:

$$G = \begin{bmatrix} g_m & 0 & \ldots & 0 \\ g_{m+1} & g_m & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g_{N-m-1} & g_{N-m-2} & \cdots & g_m \end{bmatrix}, \qquad (8)$$

g_i being the ith Markov parameter of $G(q)$. The controlled error vector is:

$$\epsilon_k = y_d - z_k = y_d - Gu_k - d_k \qquad (9)$$

and the measured error vector:

$$e_k = y_d - y_k = \epsilon_k - n_k. \qquad (10)$$

2.2 Assumptions

The disturbances d_k and n_k are assumed to be zero-mean, weakly stationary random vectors with covariance matrices R_d and R_n, respectively, and bounded higher order moments. Additionally, they are considered uncorrelated with each other. Moreover, realisations of d_k and n_k are mutually independent between repetitions.

2.3 Definitions

In this paper, a non-symmetric, real matrix X is called positive definite $X > 0$ if and only if its symmetric part $(X + X^T)/2$ is positive definite.

3. ZERO MEAN ERROR ILC

The ideal aim of ILC is to achieve a zero controlled error. When there are stochastic disturbances present in the system this objective is, unfortunately, not possible. A reasonable aim, in this case, is to make its expected value equal to zero. The goal of Zero Mean Error ILC (ZME-ILC) is, thus, to iteratively calculate the optimal input signal u^* such that:

$$E\{\epsilon(u^*)\} = E\{e(u^*)\} = 0 \qquad (11)$$

where $E\{\cdot\}$ denotes the mathematical expectation and $\epsilon(u)$ and $e(u)$ explicitly state that the error signals are functions of u.

It is straightforward to see that the solution to criterion (11) is:

$$u^* = G^{-1}y_d. \qquad (12)$$

To calculate u^* exact knowledge of G is needed, which is not possible. Nevertheless, u^* can be found using an iterative stochastic approximation (SA) procedure, such as the Robbins-Monro algorithm (Robbins and Monro, 1951), which proposes:

$$u_{k+1} = u_k + \gamma_k e_k. \qquad (13)$$

This algorithm has the form of a standard ILC law, but uses an iteration varying gain γ_k.

3.1 Almost Sure Convergence

Certain conditions must be fulfilled in order that the algorithm converges:

Theorem 1. Under the assumptions of Subsection 2.2, the iterative update algorithm (13) converges almost surely to the solution u^* of (11) when $k \to \infty$, provided that:

(C1) The sequence γ_k of positive steps satisfies:

$$\sum_{k=1}^{\infty} \gamma_k = \infty \quad \text{and} \quad \sum_{k=1}^{\infty} \gamma_k^2 < \infty. \qquad (14)$$

(C2) $E\{e(u)\}$ is monotonically decreasing:

$$Q(u) = \frac{d}{du} E\{e(u)\} < 0. \qquad (15)$$

Proof. The proof is similar to that of the Robbins-Monro stochastic approximation algorithm.

Condition **(C1)** should be fulfilled by an appropriate choice of the sequence γ_k. Condition **(C2)** is thus the primary requirement for convergence. $Q(u)$ can be rewritten as:

$$\begin{aligned} Q(u) &= \frac{d}{du} E\{e(u)\} \\ &= \frac{d}{du} E\{y_d - (Gu + d + v)\} \\ &= -G \end{aligned} \qquad (16)$$

and so condition **(C2)** requires:

$$G > 0. \qquad (17)$$

The condition (17), however, restricts the use of the algorithm to a relatively small class of systems. It is, nonetheless, possible to define the alternative criterion:

$$E\{L\epsilon(u^*)\} = E\{Le(u^*)\} = 0, \qquad (18)$$

where L is a non-singular matrix, which has the same solution (12) as criterion (11). The iterative procedure used to minimise this new criterion is:

$$u_{k+1} = u_k + \gamma_k L e_k, \qquad (19)$$

and the condition corresponding to (17) is:

$$LG > 0. \qquad (20)$$

This means that even if the system is such that G is not positive definite, an L can be sought which will make their product positive definite.

Criterion (18) will be considered for the rest of this section.

Remarks:

1) An obvious choice of L to satisfy condition (20) is $L = G^{-1}$. In certain cases, however, G can be badly conditioned so it is undesirable to use its inverse. In general, condition (20) can be seen to be a Linear Matrix Inequality (LMI) and, thus, a feasible solution L, other than the exact inverse, could be calculated using standard LMI solvers.

2) More intuitively, if L is the matrix representation of a causal, LTI filter $L(q)$, it is shown in Hatonen (2004) that a sufficient condition for (20) is that $L(q)q^m G(q)$ be strictly positive real (SPR) i.e.

$$Re[L(e^{j\omega h})e^{jm\omega h}G(e^{j\omega h})] > 0 \quad \text{for } \omega \in [0, \omega_N], \qquad (21)$$

where ω_N is the Nyquist frequency and h the sampling period. Thus, by a judicious choice of $L(q)$, condition (20) can be satisfied, even when $q^m G(q)$ is not strictly positive real.

3) The selection of L appears to require exact knowledge of G, which is never obtainable. Nonetheless, convergence of (19), using an L chosen based on a model of G, can still be assured despite a certain amount of uncertainty in the model. This issue will be examined in a later section.

3.2 Monotonic Convergence

Whilst almost sure convergence of the input sequence to the solution u^* when $k \to \infty$ is, obviously, of utmost importance, practically it is not the only type of convergence of interest. The monotonic convergence, from one iteration to the next, of a norm of the controlled error is also of concern.

By combining equations (9), (10) and (13) we can obtain the error evolution as:

$$\epsilon_{k+1} = (I - \gamma_k GL)\epsilon_k + d_k - d_{k+1} + \gamma_k GL n_k. \qquad (22)$$

We see that ϵ_{k+1} is now a function of d_k and n_k due to them being fed back into the system from the previous iteration. Taking the expected value of equation (22) we get:

$$E\{\epsilon_{k+1}\} = (I - \gamma_k GL)E\{\epsilon_k\}, \qquad (23)$$

where the zero mean assumptions on the disturbance signals have been used.

For convergence of the expected value of the controlled error to zero in 1 iteration we can see that the optimal choice of design variables is $L = G^{-1}$ and $\gamma_k = 1$. However, as already remarked, even if G were to be exactly known, it may not be desirable to use its inverse. Nevertheless, even without this choice of variables, the 2-norm of the expected value of the controlled error converges monotonically if the following condition is satisfied (see e.g. Theorem 2, Norrlöf and Gunnarsson (2002)):

$$\overline{\sigma}(I - \gamma_k GL) < 1 \quad \forall k \qquad (24)$$

where $\overline{\sigma}$ is the maximum singular value.

Lemma 2. If GL is positive definite, then condition (24) can be satisfied by making γ_0 sufficiently small.

Proof. This follows as (24) implies:

$$\overline{\lambda}([I - \gamma_k GL]^T [I - \gamma_k GL]) < 1 \quad \forall k, \qquad (25)$$

where $\overline{\lambda}$ is the maximum eigenvalue. This, in turn, gives:

$$\begin{aligned} &\overline{\lambda}(I - 2\gamma_k GL + \gamma_k^2 (GL)^T (GL)) < 1 \quad \forall k \\ \Rightarrow &1 - \underline{\lambda}(2\gamma_k GL - \gamma_k^2 (GL)^T (GL)) < 1 \quad \forall k \\ \Rightarrow &\underline{\lambda}(2GL - \gamma_k (GL)^T (GL)) > 0 \quad \forall k, \qquad (26) \end{aligned}$$

where $\underline{\lambda}$ is the minimum eigenvalue. Furthermore the eigenvalues are given by:

$$[2GL - \gamma_k (GL)^T (GL)]x_i = \lambda_i x_i, \qquad (27)$$

where x_i are the eigenvectors corresponding to the eigenvalues λ_i. Left multiplying by x_i^T we get:

$$x_i^T [2GL - \gamma_k (GL)^T (GL)]x_i = \lambda_i x_i^T x_i. \qquad (28)$$

So if GL is positive definite then in order for all λ_i to be positive, as required by (26), γ_k should satisfy:

$$2x_i^T GL x_i - \gamma_k x_i^T (GL)^T (GL)x_i > 0, \qquad (29)$$

1480

which can be achieved by choosing γ_k sufficiently small.

Additionally in order to satisfy condition **(C1)** the sequence γ_k should decrease from one iteration to the next. This means that $\gamma_0 > \gamma_k \quad \forall k > 0$. Thus if condition (29) is satisfied for $k = 0$ it will be satisfied for all other k.

Remarks:

1) It should be noted that the requirement of Lemma 2 that GL be positive definite is satisfied when L and G commute and condition (20) is satisfied.

2) When L represents a causal LTI filter, a frequency domain insight is given by noting that a sufficient condition for (24) to be fulfilled is the following (see e.g. Theorem 8, Norrlöf and Gunnarsson (2002)):

$$\max_{\omega \in [0,\omega_N]} \left| 1 - \gamma_k e^{jm\omega h} G(e^{j\omega h}) L(e^{j\omega h}) \right| < 1 \quad \forall k. \tag{30}$$

Condition (30) can be interpreted as requiring the Nyquist plot of $e^{jm\omega h} G(e^{j\omega h}) L(e^{j\omega h})$ to remain within the circle of radius $1/\gamma_k$ centred at the point $(1/\gamma_k,0)$. This frequency domain condition is complementary to that of (21). Since $L(q)$ is considered causal and thus commutes with $e^{jm\omega h} G(q)$, once condition (21) has been satisfied, γ_k can be chosen to satisfy (30). As before, this can be done by letting $\gamma_k \to 0^+$. Again as the sequence γ_k should decrease from one iteration to the next, it is only necessary to satisfy condition (30) at $k = 0$.

3.3 Robustness

As stated before, the true system G is never known exactly in practice. It is therefore important to know the conditions on the modelling uncertainty that are allowable in order that the algorithm will converge when used on the real system.

We have that the real system can be represented as:

$$G(q) = \hat{G}(q)[1 + \Delta(q)] \tag{31}$$

where $\hat{G}(q)$ is a model of the system with relative degree $\hat{m}$ and $\Delta(q)$ represents the multiplicative uncertainty. This representation is given in lifted-system form as:

$$G = \hat{G}[I + \Delta] \tag{32}$$

Substituting equation (32) into condition (20) we find:

$$L\hat{G}[I + \Delta] > 0. \tag{33}$$

In the design of the algorithm we calculate L to assure $L\hat{G} > 0$, however, in order to obtain almost sure converge of the input on the real system condition (33) should be satisfied.

Remarks:

1) For the specific choice of $L = \hat{G}^{-1}$ condition (33) becomes:

$$I + \Delta > 0. \tag{34}$$

A sufficient condition for (34) is that the filter $q^{m-\hat{m}}[1 + \Delta(q)]$ is SPR. So in the case that $m = \hat{m}$, almost sure convergence of (19) is attained when $\|\Delta(q)\|_\infty < 1$. This condition occurs frequently in the model uncertainty representation. Similar results are found in Harte et al. (2005) for this case.

2) If condition (33) is fulfilled, $m \geq \hat{m}$, which implies that Δ is lower triangular Toeplitz, and also L is lower triangular Toeplitz then $LG = GL$ is positive definite. In this case, the monotonic convergence condition (24) can still be satisfied by letting $\gamma_0 \to 0^+$. This follows from Lemma 2.

3.4 Asymptotic Distribution

The asymptotic distribution of the input estimates can be found according to the following theorem:

Theorem 3. Assume that:

i) Algorithm (19) converges almost surely to the solution u^* as $k \to \infty$.

ii) The sequence of step sizes is chosen as $\gamma_k = \frac{\alpha}{k+1}$.

iii) The matrix $D = I/2 + \alpha Q(u^*)$ has all eigenvalues with negative real parts.

Then the sequence $\sqrt{k}(u_k - u^*) \in \text{As } \mathcal{N}(0,V)$ i.e it converges asymptotically in distribution to a zero-mean normal distribution with covariance

$$V = \alpha^2 \int_0^\infty \exp(Dx) P \exp(D^T x) dx \tag{35}$$

where P is the covariance matrix of $Le(u^*)$:

$$P = E\{Le(u^*)(Le(u^*))^T\} \tag{36}$$

Proof. The proof can be found in Nevelson and Hasminskii (1973) (Theorem 6.1 p.147).

Using Theorem 3 we have that:

$$\begin{aligned} P &= E\{Le(u^*)(Le(u^*))^T\} \\ &= E\{(-L(d+n))(-L(d+n))^T\} \\ &= L(R_d + R_n)L^T, \end{aligned} \tag{37}$$

where the fact that d_k and n_k are uncorrelated has been used. Also, as $Q(u^*) = -LG$, we have that:

$$D = (I/2 - \alpha LG). \tag{38}$$

The covariance matrix V is then the unique symmetric solution of the following Lyapunov equation:

$$2\alpha^2 L(R_d + R_n)L^T + (I - 2\alpha LG)V + V(I - 2\alpha LG)^T = 0 \tag{39}$$

It is shown in Benveniste et al. (1990) (Proposition 4, p.112) that if instead of taking α as a scalar it is taken as a non-singular matrix $\alpha = K$, then the optimal gain matrix K^* to mimimise the trace of V is given by

$$K^* = -Q(u^*)^{-1} = (LG)^{-1}. \tag{40}$$

Using this gain matrix results in the learning law:

$$u_{k+1} = u_k + \frac{G^{-1}}{k+1} e_k, \tag{41}$$

and the optimal asymptotic covariance matrix:

$$V^* = G^{-1}(R_d + R_n)G^{-T}. \tag{42}$$

Thus with this optimal asymptotic covariance matrix, we have that the sequence $\sqrt{k}(u_k - u^*) \in \text{As } \mathcal{N}(0,V^*)$.

Moreover we have that $\epsilon_k = -G(u_k - u^*) - d_k(t)$ so the covariance matrix of ϵ_k is then given by:

$$\begin{aligned} \text{cov}(\epsilon_k) &= E\{\epsilon_k \epsilon_k^T\} \\ &= GE\{(u_k - u^*)(u_k - u^*)^T\}G^T + R_d. \end{aligned} \tag{43}$$

So using the optimal gain matrix K^* means that the sequence ϵ_k will have a converged covariance matrix given by $\operatorname{cov}(\epsilon_k) = \frac{1}{k}(R_d + R_n) + R_d$ and in the limit we have:

$$\lim_{k \to \infty} \operatorname{cov}(\epsilon_k) = R_d. \tag{44}$$

K^* is, however, not implementable because exact knowledge of G is not achievable. Nonetheless it gives an ideal case to aim for in the design of a stochastic ILC algorithm.

4. MINIMISATION OF THE MEAN 2-NORM OF THE CONTROLLED ERROR

A logical alternative objective to criterion (11) for an ILC algorithm is to minimise the mean 2-norm of the controlled error. In Minimal Error Norm ILC (MEN-ILC) the criterion to be minimised is thus:

$$J(u) = \frac{1}{2} E\{\epsilon(u)^T \epsilon(u)\}. \tag{45}$$

The minimum of this criterion is found when:

$$
\begin{aligned}
J'(u^*) &= \left. \frac{dJ(u)}{du} \right|_{u^*} \\
&= E\left\{ \left(\left. \frac{\partial \epsilon(u)}{\partial u} \right|_{u^*} \right)^T \epsilon(u^*) \right\} \\
&= -G^T E\{\epsilon(u^*)\} = 0, \tag{46}
\end{aligned}
$$

which clearly holds when $u^* = G^{-1} y_d$, i.e. equation (12) again. However, as before, exact knowledge of G is not realistic, and it is not possible to measure ϵ directly in order to calculate the optimal input iteratively using it. Nonetheless, because equation (46) can be written as

$$
\begin{aligned}
J'(u^*) &= -G^T E\{\epsilon(u^*)\} \\
&= -G^T E\{e(u^*)\} = 0 \tag{47}
\end{aligned}
$$

it is possible to find the minimum of the criterion, again, using the Robbins-Monro algorithm:

$$u_{k+1} = u_k + \gamma_k G^T e_k. \tag{48}$$

4.1 Convergence

Comparing (48) with (19) we see that $L = G^T$. Thus the almost sure convergence condition (20) is automatically satisfied, as G is non-singular.

Additionally, monotonic convergence is assured with a sufficiently small choice of γ_k. This property follows from Lemma 2 and the fact that $GL = GG^T$ is positive definite.

4.2 Estimating the gradient

The algorithm converges to the minimum of (45) provided that an unbiased estimate of $J'(u_k)$, or alternatively $G^T e$, is used in (48).

a) Use of a model: An estimate of $G^T e$ can be achieved by replacing G^T with $\hat{G}^T$, the model. The model will have uncertainty, however.

With $L = \hat{G}^T$, condition (33) gives:

$$\hat{G}^T \hat{G}[I + \Delta] > 0. \tag{49}$$

If the assumption that the relative degree of $G(q)$ is greater than or equal to $\hat{G}(q)$ is made, then $I + \Delta$ is a lower triangular Toeplitz matrix. $I + \Delta$ therefore commutes with

$\hat{G}$ and (49) can be written $\hat{G}^T [I + \Delta] \hat{G} > 0$. This condition is fulfilled when $\hat{G}$ is non-singular and $I + \Delta > 0$. $\hat{G}$ is always nonsingular because N is finite, thus it is necessary only that $I + \Delta > 0$ i.e. the same condition as found in Subsection 3.3 when $L = \hat{G}^{-1}$. This means that a certain amount of uncertainty in the model can be accepted and the algorithm will still almost surely converge to the minimum of (45).

a) Use of an experiment: Alternatively, by noting that $e2 = G^T e_k$ can be interpreted as the following filtering operations:

$$e1(t) = G(q)e_k(N - t) \tag{50}$$
$$e2(t) = e1(N - t), \tag{51}$$

we see that, in the disturbance free case, the gradient can be found by an experiment on the true system, where the time reversed error signal is fed into the system as its input and the system output is measured and then time reversed itself before being used as $G^T e = e2$. In reality the special experiment will have its own disturbances $d2(t)$ and $v2(t)$ associated with it. However, an unbiased estimate of $G^T e$ can still be found as:

$$
\begin{aligned}
E\{e2\} &= E\{G^T e_k + d2 + v2\} \\
&= E\{G^T e_k\} + E\{d2\} + E\{n2\} \tag{52} \\
&= G^T E\{\epsilon_k\} + 0 + 0. \tag{53}
\end{aligned}
$$

This method of estimating the gradient is attractive as avoids the problems of model uncertainty. It does, however, require an additional, non-standard, experiment at each iteration.

4.3 Connections to previously proposed algorithms

The proposed model free algorithm has similarities to that proposed in Ye and Wang (2005) where reversed time inputs are used to cancel the system phase and produce monotonic convergence, but there stochastic aspects are not considered.

It also has similarities to Hatonen et al. (2003), which as mentioned beforehand uses the steepest descent method, and calls G^T the adjoint of G. It shows that by using this 'adjoint' with an iteration-varying gain, monotonic convergence occurs. The gain sequence is calculated via an optimisation, which does not consider stochastic disturbances. The gain at iteration k is given by:

$$\gamma_k = \frac{\|G^T e_{k-1}\|^2}{w + \|GG^T e_{k-1}\|^2}, \tag{54}$$

where w is a weight on γ_k in the cost function. Since the measured error signal is used to calculate the gain, it will be affected by stochastic disturbances. This means $\lim_{k \to \infty} \|G^T e_{k-1}\|^2 \neq 0$ and so $\lim_{k \to \infty} \gamma_k \neq 0$. This means that the second series of condition **C1** cannot be satisfied. Therefore, whilst the algorithm developed can lead to fast deterministic convergence to the optimal input, it is not possible to prove this when stochastic disturbances are present.

5. CONCLUSIONS

Stochastic approximation theory has been applied to Iterative Learning Control to produce algorithms which are ro-

bust to non-repetitive disturbances, a blight for the learning process of standard ILC algorithms. Two algorithms have been proposed, one to obtain zero mean controlled error (ZME-ILC) and one to minimise the mean 2-norm of the error (MEN-ILC). ZME-ILC requires knowledge of the system, to within certain uncertainty bounds, but in the presence of noise can give reasonably rapid convergence. MEN-ILC can either use a model or be model free, though it then requires twice as many experiments to be carried out on the system. Furthermore MEN-ILC can give good tracking improvement for non-minimum phase systems.

It is noted that the algorithms proposed require the learning gain to tend to zero as the iterations tend to infinity. This requirement seems to be essential for stochastic learning algorithms. Practically it means the learning ceases after a large number of iterations and if the desired output or repetitive disturbances change the algorithm will not react and the tracking will deteriorate. It is thus necessary to have a surveillance program that restarts the learning when the errors rise above a certain threshold.

REFERENCES

H. S. Ahn, K. L. Moore, and Y. Chen. Kalman filter augmented iterative learning control on the iteration domain. In *IEEE American Control Conference*, pages 250–255, Minneapolis, U.S.A., 2006.

S. Arimoto. Robustness of learning control for robotic manipulators. In *IEEE International Conference on Robotics and Automation*, pages 1528–1533, Cincinnati, Ohio USA, May 1990.

A. Benveniste, M. Metivier, and P. Priouret. *Adaptive Algorithms and Stochastic Approximation*. Springer-Verlag, Berlin and New York, 1990.

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

H. F. Chen. Almost sure convergence of iterative learning control for stochastic systems. *Sci. in China (Series F)*, 46(1):1–13, 2003.

H. F. Chen and H. T. Fang. Output tracking for nonlinear stochastic systems by iterative learning control. *IEEE Transactions on Automatic Control*, 49(4):583–588, 2004.

S. Gunnarsson and M. Norrlöf. On the disturbance properties of high order iterative learning control algorithms. *Automatica*, 42(11):2031–2034, 2006.

T. J. Harte, J. Hatonen, and D. H. Owens. Discrete-time inverse model-based iterative learning control: stability, monotonicity and robustness. *International Journal of Control*, 78(8):577–586, May 2005.

J. Hatonen. *Issues of algebra and optimality in iterative learning control*. PhD thesis, University of Oulu, Finland, 2004.

J. J. Hatonen, T. J. Harte, D. H. Owens, J. D. Radcliffe, P. L. Lewin, and E. Rogers. A new robust iterative learning control algorithm for application on a gantry robot. In *IEEE Conference on Emerging Technologies in Factory Automation*, pages 305–312, Lisbon, Portugal, 2003.

G. Heinzinger, D. Fenwick, B.Paden, and F. Miyaziki. Robust learning control. In *28th IEEE Conference on Decision and Control*, pages 436–440, Tampa, Florida USA, December 1989.

M. B. Nevelson and R. Z. Hasminskii. *Stochastic Approximation and Recursive Estimation*. American Mathematical Society, Providence, Rhode Island, USA, 1973.

M. Norrlöf. An adaptive iterative learning control algorithm with experiments on an industrial robot. *IEEE Transactions on Robotics and Automation*, 18(2):245–251, April 2002.

M. Norrlöf and S. Gunnarsson. Disturbance aspects of iterative learning control. *Engineering Applications of Artificial Intelligence*, 14(1):87–94, 2001.

M. Norrlöf and S. Gunnarsson. Time and frequency domain convergence properties in iterative learning control. *International Journal of Control*, 75(14):1114–1126, 2002.

H. Robbins and S. Monro. A stochastic approximation method. *Ann. Math. Stat.*, 22:400–407, 1951.

S. S. Saab. Optimal selection of the forgetting matrix into an iterative learning control algorithm. *IEEE Transactions on Automatic Control*, 50(12):2039–2043, December 2005.

S.S. Saab. A discrete-time stochastic learning control algorithm. *IEEE Transactions on Automatic Control*, 46(6):877–887, June 2001a.

S.S. Saab. On a discrete-time stochastic learning control algorithm. *IEEE Transactions on Automatic Control*, 46(8):1333–1336, August 2001b.

M.K. Tao, R.L. Kosut, and A. Gurcan. Learning feedforward control. In *IEEE American Control Conference*, pages 2575–2579, Baltimore, Maryland USA, June 1994.

Y. Ye and D. Wang. Zero phase learning control using reversed time input runs. *Journal of Dynamic Systems, Measurement and Control*, 127:133–139, March 2005.