Concurrent Learning Adaptive Control with Directional Forgetting

Hae-In Lee, Hyo-Sang Shin, Antonios Tsourdos

Abstract—This paper proposes a new concurrent learning based adaptive control algorithm. The main objective behind our proposition is to relax the persistent excitation requirement for the stability guarantee, while providing the ability to identify time-varying parameters. To achieve the objective, this paper designs a directional forgetting algorithm, which is then integrated with the adaptive law. The theoretical stability analysis shows that the tracking and parameter estimation error is exponentially stable with the signal only finitely excited, not persistently excited. The analysis also shows that the proposed algorithm can guarantee the stability under time-varying parameters. Moreover, the necessary and sufficient conditions for the stability given the time-varying parameters are derived. The results of numerical simulations confirm the validity of the theoretical analysis results and demonstrate the performance of the proposed algorithm.

Index Terms—Directional Forgetting, Model Reference Adaptive Control, Concurrent Learning

I. INTRODUCTION

THE model reference adaptive control (MRAC) has been widely used for estimating the uncertainty and canceling out its effect from the system to achieve the nominal designed performance [1]. One of the main issues of the MRAC is that the persistence of excitation (PE) is required for the parameter estimation to be converged. The PE corresponds to the continuous change in the states, which is undesirable for the control performance as it contradicts with obtaining the steady-state, and may contribute to the waste of energy.

There have been great efforts made to relax the PE requirement. Earlier studies on the relaxation of the PE requirement for parameter convergence include the concurrent learning (CL) based method [2], [3] and its modifications [4]–[6]. Conceptually, the PE condition is required since the adaptive law is rank-1 and thus the inputs are required to be persistently excited in every direction to span the parameter estimation error. The common principle to relax the dependence on PE in CL based methods is to use the stored data containing information from the Finite Excitation (FE) in the past together with the current data. In this way, the rank deficiency of an information matrix is expected to be solved after sufficient accumulation of data, and this is the key of relaxation.

The issue with these methods is that the parameter convergence is difficult to be guaranteed if parameters are time varying. The information accumulated before a parameter change contains only the information about the previous parameters. After the parameter change, the information about

H.-I. Lee, H.-S. Shin, and A. Tsourdos are with the School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, United Kingdom. e-mail: (haein.lee ; h.shin ; a.tsourdos@cranfield.ac.uk)

the parameters changed starts to be accumulated. Roughly speaking, since the information matrix contains both the previous parameters and those changed, utilising this information matrix is difficult to guarantee the convergence to time-varying parameters. Especially when the information about previous parameters is rich, the convergence issue could be exacerbated. The CL based methods thus may not guarantee the stability of parameter estimation for the systems with time-varying parameters, such as the control of strip temperature for heating furnaces, automation of the heavy duty vehicles, and self-tuning cruise control [7]–[9].

Forgetting the previous information could address convergence issue under the time-varying parameters: various forgetting algorithms have been used in the online parameter estimation [10] to cope with the time-varying parameters. Note that an adaptive control architecture for uncertain dynamic system consists of two principal components: one is the adaptive element in the control law and the other is the adaptation law. The adaptive element in the control law is usually a function approximator to cancel out the effect of uncertainty from the tracking error dynamics. The adaptation law is in principle a regression algorithm working for better approximation of uncertainty. Therefore, it is clear that the forgetting methods can be applied to the adaptation law and thus integrated to CL methods.

Cho *et al.* [6] developed a composite MRAC algorithm to relax the PE requirement. Based on the filtered regressor, they integrated the exponential forgetting method in the composite MRAC. The potential issue is that their proposed algorithm requires PE condition for the forgetting algorithm. To avoid the PE condition, which is required for the stability guarantee in other methods with either forgetting [11], [12] or finite-time identifier [13] for the stability guarantee, they applied the exponential forgetting only when the minimum eigenvalue of the stacked data increases. However, the forgetting is not working in the modified forgetting algorithm when the signal is not persistently exited. Therefore, the PE condition is again crucial either for the convergence of parameter estimation or for the forgetting of past information.

This paper proposes a new CL adaptive control algorithm to relax the PE requirement and also to handle the convergence issue under time-varying parameters. The key idea of the proposed adaptive control algorithm is to integrate directional forgetting to the adaptation law. Note that the directional forgetting method [14]–[16] was intended to avoid the estimator windup – zero eigenvalue of the information matrix – by forgetting the old data only in the direction of new data. In this paper, its objective is extended to guarantee

the stability of the both tracking and parameter estimation error. To achieve this objective, we modify the directional forgetting proposed in [16] and apply it to the adaptation law. As the modified directional forgetting accumulates the data and maintains the stacked data, the PE condition can be directly relaxed. Moreover, since it discounts the information in the direction of new data, forgetting is always working on the non-zero signals after FE. This could relax the PE condition for forgetting the past data and consequently provide the convergence of time-varying parameters.

The characteristics of the new CL adaptive are analytically and empirically investigated. The stacked data is first proven to be always lower bounded by a positive value and thus full-rank. Then, this paper shows that the past information is consistently forgotten but maintains its ability to estimate parameter uncertainty without requiring the PE condition. The stability of tracking and parameter estimation under the presence of parameter change is also theoretically analysed by assuming the discrete changes in parameters. The stability conditions depend on the size of parameter change and forgetting rate, providing the design trade-offs of forgetting rate. The numerical results on the wingrock dynamics confirm the stability analysis results and demonstrate the effect of forgetting rate.

The rest of the paper is organised as follows. In Section II, mathematical preliminaries with definitions and lemmas are given for later proofs. The control problem with the parameterised uncertainty is formulated in Section III. In Section IV, the adaptive control with directional forgetting is suggested, and the bounds of the information and the stability conditions are examined. The numerical simulations in Section V show the performance of the proposed control and support the theoretic stability conditions. The paper is concluded in Section VI.

II. PRELIMINARIES

The PE condition is crucial for the common MRAC methods and the parameter estimation with exponential forgetting. This paper intends to relax the PE condition to FE which requires exciting signals only for a finite time interval. The PE and FE conditions are defined as in [17].

Definition 1 (Persistence of Excitation (PE)). *A bounded vector signal $q(t)$ is persistently exciting if there exist $T > 0$ and $\gamma > 0$ such that*

$$\int_t^{t+T} q(\tau)q(\tau)^T d\tau \geq \gamma I, \qquad \forall t \geq t_0. \tag{1}$$

Definition 2 (Finite Excitation (FE)). *A bounded vector signal $q(t)$ is finitely exciting over a time set $[t_s, t_s+T]$ if there exist $t_s \geq t_0$, $T > 0$ and $\gamma > 0$ such that*

$$\int_{t_s}^{t_s+T} q(\tau)q(\tau)^T d\tau \geq \gamma I. \tag{2}$$

The following Lemma utilises the characteristics of spectral radius to obtain the bound of a matrix. This Lemma is used in Theorem 1 to prove the lower bound of the information.

Lemma 1. *For a positive semi-definite matrix $A \in \mathbb{R}^{n \times n}$, $A \leq I$ if and only if $\rho(A) \leq 1$, where the spectral radius $\rho(\cdot)$ is defined as the largest absolute value of the eigenvalues of a matrix.*

III. PROBLEM FORMULATION

A. System Dynamics

Consider a state-space representation as:

$$\dot{x}(t) = Ax(t) + B(u(t) + \Delta(t)), \tag{3}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, and $\Delta(t) \in \mathbb{R}^m$ stand for state, input, and uncertainty vector, respectively. The system matrices $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$ are assumed to be constant and controllable.

The uncertainty $\Delta(t)$ is assumed to be linearly parametrised as:

$$\Delta(t) = W^{*T}(t)\Phi(x(t)), \tag{4}$$

where $W^*(t) \in \mathbb{R}^{l \times m}$ is the unknown true parameter matrix, and $\Phi(x(t)) \in \mathbb{R}^l$ is the basis vector.

B. Baseline Controller and Tracking Error Dynamics

The controller is designed in two parts as:

$$u(t) = u_{base}(t) - u_{ad}(t), \tag{5}$$

where the baseline control $u_{base}(t)$ determines the nominal performance of the control and the adaptive control $u_{ad}(t)$ alleviates the effect of uncertainty.

The baseline controller is designed as $u_{base}(t) = -Kx(t) + K_r r(t)$ such that the system is stable and tracks the reference input $r(t) \in \mathbb{R}^m$. The reference model, the model without any parameter uncertainty, is obtained as:

$$\dot{x}_r = A_r x_r(t) + B_r r(t) \tag{6}$$

where $x_r(t) \in \mathbb{R}^n$ is the reference state. The input matrix of the reference model is defined with respect to the reference input matrix $K_r \in \mathbb{R}^{m \times m}$ as $B_r = BK_r$. The control gain $K \in \mathbb{R}^{m \times n}$ is determined such that the system matrix of the reference model, $A_r = A - BK$, is Hurwitz stable. Then, there exists a positive definite symmetric matrix $P \in \mathbb{R}^{n \times n}$ satisfying the following Lyapunov equation:

$$A_r^T P + PA_r = Q. \tag{7}$$

If there exists the matrix K_r such that $B_r = -A_r$, the reference state $x_r(t)$ converges to the reference input $r(t)$. Here, the adaptive control is designed to reduce the error between the reference and real model, and to eventually obtain the states converge to the reference input $r(t)$. Defining the tracking error as $e(t) = x_r(t) - x(t)$, the tracking error dynamics is given by

$$\dot{e}(t) = A_r e(t) + B\epsilon(t), \tag{8}$$

where $\epsilon(t) \triangleq u_{ad}(t) - \Delta(t)$ is the adaptation error.

Assuming that the uncertainty lies in the span of the input matrix B, the uncertainty is accessed as:

$$\Delta(t) = B^+(\dot{e}(t) - A_r e(t)) + u_{ad}(t), \tag{9}$$

where B^+ is a pseudoinverse of B.

Note that the derivative of the tracking error $\dot{e}(t)$ cannot be perfectly known and should be obtained based on measured signals which contain measurement noise. To alleviate this issue, various filters, such as fixed-point smoother [4] and a low pass filter [6], can be used. Using one of the filters, the filtered plane of the equation Eqn. (9) is expressed as:

$$c(t) = W^{*T}(t)q(t), \qquad (10)$$

where $c(t)$ and $q(t)$ denote for the filtered vectors of $\Delta(t)$ and $\Phi(z(t))$, respectively. The filtered parameter is assumed the same as $W^*(t)$. Here, $\Delta(t)$ is obtained by approximating Eqn. (9) at every time step, and $\Phi(z(t))$ is the basis vector with respect to $z(t)$. For the details on approximating $\Delta(t)$ in Eqn. (9), please refer to [6].

IV. ADAPTIVE CONTROL

A. Adaptive Law

The adaptive control is designed to cancel out the effect of uncertainty from the tracking error dynamics in Eqn. (8) by estimating the parameter as:

$$u_{ad}(t) = \hat{W}^T(t)\Phi(z(t)). \qquad (11)$$

Here, the estimated parameter vector $\hat{W}(t) \in \mathbb{R}^{l \times m}$ is determined by an adaptive law. The adaptive law used in this paper is the summation of gradient descent method and the information-based parameter estimation:

$$\dot{\hat{W}}(t) = \Gamma\left(\Phi(z(t))e^T(t)PB - \gamma_b(\Omega(t)\hat{W}(t) - M(t))\right), \qquad (12)$$

where $\Gamma \in \mathbb{R}^{l \times l}$, and $\gamma_b \in \mathbb{R}^{l \times l}$ are the adaptive gains for gradient descent and information architecture, respectively. Also, $\Omega(t) \in \mathbb{R}^{l \times l}$, and $M(t) \in \mathbb{R}^{l \times l}$ are the information matrix and auxiliary matrix, respectively. The gradient descent terms directly reduce the tracking error, and information architecture stores the past data to maintain the convergence of parameter estimation without requiring the PE condition.

The information matrix is the accumulation of measured basis vectors, and the auxiliary matrix is that of the filtered uncertainty. Once the FE condition is satisfied, both the information and auxiliary matrices are forgotten with the directional forgetting method. The dynamics of the accumulation and forgetting method is expressed as

$$\dot{\Omega}(t)$$
$$= \begin{cases} q(t)q^T(t) & \text{if } rank(\Omega(t)) < rank(\Omega(t) + q(t)q^T(t)) \\ -k\dfrac{\Omega(t)q(t)q^T(t)}{q^T(t)\Omega(t)q(t)}\Omega(t) + q(t)q^T(t) & \text{otherwise} \end{cases}$$
$$\dot{M}(t)$$
$$= \begin{cases} q(t)c^T(t) & \text{if } rank(\Omega(t)) < rank(\Omega(t) + q(t)q^T(t)) \\ -k\dfrac{\Omega(t)q(t)q^T(t)}{q^T(t)\Omega(t)q(t)}M(t) + q(t)c^T(t) & \text{otherwise} \end{cases}$$
$$(13)$$

where $k \in \mathbb{R}$ is a positive constant. The initial values of the information and auxiliary matrices are zero.

The physical meaning of the dynamics of forgetting is that the information is forgotten only on the direction that the new

$q(t)$ is added. For example, assume that the two independent vectors $q_1 \in \mathbb{R}^2$ and $q_2 \in \mathbb{R}^2$ are stacked in the information matrix, such that $\Omega(t) = q_1q_1^T + q_2q_2^T$ is full-rank. Then, the derivative of the information matrix with the new input $q(t) \perp q_2(t)$ satisfies

$$\dot{\Omega}(t)q = -k\frac{(q_1q_1^T + q_2q_2^T)qq^T}{q^T(q_1q_1^T + q_2q_2^T)q}(q_1q_1^T + q_2q_2^T) + qq^T$$
$$= -k\frac{(q_1q_1^T)qq^T}{q^T(q_1q_1^T)q}(q_1q_1^T) + qq^T. \qquad (14)$$

From the equation, it can be implied that the stacked input in the direction orthogonal to $q(t)$ is not forgotten, whereas the others are forgotten exponentially.

B. Lower and Upper Bounds of the Information Matrix

The value of the information matrix is directly related to the convergence rate of the parameter estimation as in Eqn. (12), and thus the bounds of the information matrix are inferred to be crucial for proving the stability of parameter estimation and for finding the convergence rate. The following theorems give the lower and upper bounds of the information matrix. Note that the upper boundedness of the information matrix is guaranteed for the bounded signal, i.e. $|q(t)| < c$, as other forgetting algorithms [11].

Theorem 1 (Lower bound of the information matrix). *If the signal is finitely exciting over $[t_0, t_1]$, there exists a constant $\alpha > 0$ such that*

$$\Omega(t) \geq \alpha I, \qquad \forall t \geq t_1. \qquad (15)$$

Proof. Let the subspace which is excited by $q(t)$ be ϕ with its dimension m, and the unexcited subspace be $\phi^\perp$. The information matrix can be decomposed as:

$$\Omega(t) = \Omega_o(t) + \Omega_p(t), \qquad \Omega_o(t)q(t) = 0, \qquad (16)$$

where $\Omega_o(t)$ is orthogonal to $q(t)$ and $\Omega_p(t)$ is the remainder to be forgotten. The dynamics of each part $\Omega_o(t)$ and $\Omega_p(t)$ is defined by following equations such that the summation of the dynamics satisfies Eqn. (13).

$$\dot{\Omega}_o(t) = 0$$
$$\dot{\Omega}_p(t) = -k\frac{\Omega_p(t)q(t)q^T(t)}{q^T(t)\Omega_p(t)q(t)}\Omega_p(t) + q(t)q^T(t) \qquad (17)$$

Defining an orthogonal matrix $U = [U_1 \quad U_2]$ where the columns of $U_1 \in \mathbb{R}^{n \times m}$ are the orthogonal basis for ϕ and those of $U_2 \in \mathbb{R}^{n \times (n-m)}$ are the orthogonal basis for $\phi^\perp$, the following equation is obtained:

$$U^T\dot{\Omega}(t)U \geq$$
$$\begin{bmatrix} \left(-kU_1^T\dfrac{\Omega_p(t)q(t)q^T(t)}{q^T(t)\Omega_p(t)q(t)}U_1\right)U_1^T\Omega_p(t)U_1 & 0 \\ +U_1^Tq(t)q^T(t)U_1 & \\ 0 & 0 \end{bmatrix} \qquad (18)$$

The forgetting part, $U_1^T \Omega_p(t) U_1$ is structured with its value at $t = t_1$ as:

$$U_1^T \Omega_p(t) U_1 = \Psi(t, t_1) U_1^T \Omega_p(t_1) U_1 \\ + \int_{t_1}^{t} \Psi(t, \tau) U_1^T q(\tau) q^T(\tau) U_1 d\tau \quad (19)$$

where $\Psi(t, t_1) \in \mathbb{R}^{l \times l}$ is the transition matrix from t_1 to t:

$$\Psi(t, t_1) = \exp\left(-k \int_{t_1}^{t} F(\tau) d\tau\right), \\ F(\tau) \triangleq U_1^T \frac{\Omega_p(\tau) q(\tau) q^T(\tau)}{q^T(\tau) \Omega_p(\tau) q(\tau)} U_1 \quad (20)$$

Here, the spectral radius of the matrix $F(\tau)$ is upper bounded by 1 as:

$$\sum_i \lambda_i(F(\tau)) = tr(F(\tau)) = 1, \quad \lambda_i(F(\tau)) \geq 0 \quad \forall i, \quad (21)$$

where $\lambda_i(\cdot)$ is the i-th eigenvalue of a matrix.

Using Lemma 1 and the upper bound of the matrix $F(\tau)$, the lower bound of the transition matrix is obtained as:

$$\Psi(t, t_1) \geq e^{-k(t-t_1)} I. \quad (22)$$

As $U_1^T q(\tau)$ acts as a persistently exciting signal, i.e. $U_1^T q(t) q^T(t) U_1 > 0$, there exists a constant $\alpha > 0$ such that

$$U_1^T \Omega_p(t) U_1 \geq e^{-k(t-t_1)} U_1^T \Omega_p(t_1) U_1 + \alpha I. \quad (23)$$

The non-forgetting part, $U_2^T \Omega_o(t) U_2$, is obtained as:

$$U_2^T \Omega_o(t) U_2 = U_2^T \Omega_o(t_1) U_2 \quad (24)$$

For $\Omega(t_1) \geq \gamma I$, the solution of the equation (18) is thus lower-bounded as:

$$U^T \Omega(t) U \geq \min(\gamma e^{-k(t-t_1)} + \alpha, \gamma) I \quad (25)$$

Therefore, the lower bound of the information matrix is obtained as:

$$\Omega(t) \geq \alpha I \quad (26)$$

$\square$

Theorem 2 (Upper bound of the information matrix)**.** *If the signal is finitely exciting over $[t_0, t_1]$ and the signal is bounded, i.e. $|q(t)| < c$, there exists a constant $\beta > 0$ such that*

$$\Omega(t) \leq \beta I, \qquad \forall t > t_1. \quad (27)$$

Proof. The forgetting dynamics of the information matrix results in

$$\Omega(t) = \Omega(t_1) + \int_{t_1}^{t} \left(q(\tau) q^T(\tau) - k \frac{\Omega(\tau) q(\tau) q^T(\tau)}{q^T(\tau) \Omega(\tau) q(\tau)} \right) d\tau \quad (28)$$

The trace of the information matrix is computed as:

$$tr(\Omega(t)) = tr(\Omega(t_1)) + \int_{t_1}^{t} \frac{q^T(\tau)(|q(\tau)|^2 I - k\Omega^2(\tau)) q(\tau)}{q^T(\tau) \Omega(\tau) q(\tau)} d\tau \quad (29)$$

Assume that the information matrix is unbounded from above, i.e. there exists an eigenvalue λ_Ω that diverges to ∞. Then, the integration term becomes negative and unbounded:

$$\lim_{\lambda_\Omega \to \infty} \int_{t_1}^{t} \frac{q^T(\tau)(|q(\tau)|^2 I - k\Omega^2(\tau)) q(\tau)}{q^T(\tau) \Omega(\tau) q(\tau)} d\tau = -\infty \quad (30)$$

With the positive and boundedness of the left-hand-side, $tr(\Omega(t))$, the equation is contradictory. Therefore, the information matrix is upper bounded by a constant:

$$\Omega(t) \leq \beta I, \qquad \forall t > t_1. \quad (31)$$

$\square$

C. Stability Analysis

Using that the information matrix is lower-bounded to be positive definite even without the persistent excitation, the tracking error and the parameter estimation error can be proved to be stable by the following theorem.

Theorem 3 (Stability in $(e, \tilde{W})$)**.** *If the signal is finitely exciting over $[t_0, t_1]$, the tracking error e and the parameter estimation error $\tilde{W}$ are globally uniformly exponentially stable for $t > t_1$.*

Proof. The explicit forms of the information matrix and the auxiliary matrix are expressed as:

$$\begin{cases} \Omega(t) = \Psi(t, t_0) \Omega(t_0) + \int_{t_0}^{t} \Psi(t, \tau) q(\tau) q^T(\tau) d\tau, \\ M(t) = \Psi(t, t_0) M(t_0) + \int_{t_0}^{t} \Psi(t, \tau) q(\tau) c^T(\tau) d\tau, \end{cases} \quad (32)$$

where the transition matrix is defined as:

$$\Psi(t, t_0) = \exp\left(-k \int_{t_0}^{t} \frac{\Omega(\tau) q(\tau) q^T(\tau)}{q^T(\tau) \Omega(\tau) q(\tau)} d\tau\right). \quad (33)$$

As the initial values of the information and auxiliary matrix are zero, the following equation is obtained:

$$\Omega(t) \hat{W}(t) - M(t) = \Omega(t) \tilde{W}(t). \quad (34)$$

The Lyapunov function is defined as:

$$V(e, \tilde{W}) = \frac{1}{2} e^T(t) P e(t) + \frac{1}{2} tr(\tilde{W}^T(t) \Gamma^{-1} \tilde{W}(t)), \quad (35)$$

where $\tilde{W}(t) \triangleq \hat{W}(t) - W^*(t)$ is the parameter estimation error. Defining an augmented vector $\xi(t) \triangleq [e^T(t), \tilde{W}^T(t)]$, the upper bound of the Lyapunov function is given as:

$$V(e, \tilde{W}) \leq \frac{1}{2} \max(\lambda(P), \lambda(\Gamma^{-1})) \|\xi(t)\|^2, \quad (36)$$

where $\lambda(\cdot)$ is the eigenvalue of a matrix.

Substituting the tracking dynamics in Eqn. (8) and the adaptive law in Eqn. (12), the derivative of the Lyapunov function is obtained as:

$$\dot{V}(e, \tilde{W}) = -\frac{1}{2} e^T(t) Q e(t) - tr(\tilde{W}^T(t) \gamma_b \Omega(t) \tilde{W}(t)). \quad (37)$$

From the Theorem 1, the information matrix is lower bounded, and thus the derivative of the Lyapunov function is upper bounded as:

$$\dot{V}(e, \tilde{W}) \leq -\frac{1}{2} \min(\lambda(Q), 2\gamma_b \alpha) \|\xi\|^2. \quad (38)$$

Using the upper bound of the Lyapunov function in Eqn. (36),

$$\dot{V}(e, \tilde{W}) \leq -\frac{\min(\lambda(Q), 2\gamma_b \alpha)}{\max(\lambda(P), \lambda(\Gamma^{-1}))} V(e, \tilde{W}). \quad (39)$$

As $V(e, \tilde{W}) > 0$ for all $(e, \tilde{W})$ except the origin, and the lower bound of the information matrix is positive, i.e. $\alpha > 0$, by Theorem 1, $(e, \tilde{W})$ are globally uniformly exponentially stable. $\qquad\square$

One of the main reasons of the forgetting algorithms is to cope with the parameter changes. Assume the parameter change with the size of ΔW_i^* at $t = t_i$ as:

$$W^*(t) = W_0^* + \sum_i \Delta W_i^* h(t - t_i), \qquad (40)$$

where $h(t)$ is the step function.

The sufficient and necessary stability condition under the presence of parameter change is derived in the following theorem.

Theorem 4 (Stability in $(e, \tilde{W})$ with parameter changes). *Given the finitely exciting signal, the tracking error e and the parameter estimation error $\tilde{W}$ are Lyapunov stable for $t \in [t_i, t_{i+1}]$ if and only if*

$$
\begin{aligned}
F(e, \tilde{W}) &\triangleq \frac{1}{2} e^T(t) Q e(t) I + \Omega(t) \tilde{W}(t) \tilde{W}^T(t) \\
&+ \sum_{j=1}^{i} \Psi(t, t_j) \Omega(t_j) \Delta W_j^* \tilde{W}^T(t) \geq 0,
\end{aligned}
\qquad (41)
$$

where the transition matrix $\Psi(t, t_j)$ is given by

$$\Psi(t, t_j) = \exp\left(-k \int_{t_j}^{t} \frac{\Omega(\tau) q(\tau) q^T(\tau)}{q^T(\tau) \Omega(\tau) q(\tau)} d\tau\right). \qquad (42)$$

Proof. For the time span $t \in [t_i \quad t_{i+1}]$, the information matrix and the auxiliary matrix are expressed as:

$$
\begin{cases}
\Omega(t) = \Psi(t, t_i) \Omega(t_i) + \int_{t_i}^{t} \Psi(t, \tau) q(\tau) q^T(\tau) d\tau \\
M(t) = \Psi(t, t_i) M(t_i) + \int_{t_i}^{t} \Psi(t, \tau) q(\tau) c^T(\tau) d\tau
\end{cases}
\qquad (43)
$$

As the values of both information and auxiliary matrices are continuous at the parameter jumps, the following equation is satisfied.

$$
\begin{aligned}
&\Omega(t) \hat{W}(t) - M(t) \\
&= \Omega(t) \tilde{W}(t) + \sum_{j=1}^{i} \Psi(t, t_j) \Omega(t_j) \Delta W_j^*
\end{aligned}
\qquad (44)
$$

Substituting the equation into the same Lyapunov function as the equation (35), the derivative of the Lyapunov function is computed as:

$$
\begin{aligned}
\dot{V}(e, \tilde{W}) &= -\frac{1}{2} e^T(t) Q e(t) - tr(\tilde{W}^T(t) \gamma_b \Omega(t) \tilde{W}(t)) \\
&- \sum_{j=1}^{i} tr(\tilde{W}^T(t) \gamma_b \Psi(t, t_j) \Omega(t_j) \Delta W_j^*)
\end{aligned}
\qquad (45)
$$

The necessary and sufficient condition for $\dot{V}(e, \tilde{W}) \leq 0$ for all $(e, \tilde{W})$ except the origin is

$$
\begin{aligned}
&\frac{1}{2} e^T(t) Q e(t) I + \Omega(t) \tilde{W}(t) \tilde{W}^T(t) \\
&+ \sum_{j=1}^{i} \Psi(t, t_j) \Omega(t_j) \Delta W_j^* \tilde{W}^T(t) \geq 0
\end{aligned}
\qquad (46)
$$

$\square$

In the Theorem 4, the sufficient and necessary condition $F(e, \tilde{W})$ is not deterministic as the transition matrix $\Psi(t, t_j)$ for the time-varying system is almost impossible to obtain. Instead, the following propositions separate the sufficient condition $F_1(e, \tilde{W})$ and the necessary condition $F_2(e, \tilde{W})$, of which the bounds are conservative, but can be computed deterministically.

Proposition 1. *The tracking error e and the parameter estimation error $\tilde{W}$ are exponentially stable for $t \in [t_i, t_{i+1}]$ if $\Delta W_j^* \tilde{W}^T(t_j) \geq 0$ for all $j \in [1, i]$.*

Proof. If $\Delta W_j^* \tilde{W}^T(t) \geq 0$, the derivative of the Lyapunov function under the presence of parameter change in the equation (45) is less than or equal to that without the parameter change in the equation (37). Converging faster than the exponentially stable case, $(e, \tilde{W})$ is also exponentially stable. $\quad\square$

Proposition 2. *The tracking error e and the parameter estimation error $\tilde{W}$ are Lyapunov stable if $F_1(e, \tilde{W}) \geq 0$ for $t \in [t_i, t_{i+1}]$, where the function $F_1(e, \tilde{W})$ is defined as:*

$$
\begin{aligned}
F_1(e, \tilde{W}) &= \frac{1}{2} e^T(t) Q e(t) I \\
&+ \alpha \tilde{W}(t) \tilde{W}^T(t) + \beta \sum_{j=1}^{i} \Delta W_j^* \tilde{W}^T(t)
\end{aligned}
\qquad (47)
$$

Proof. If $\Delta W_j^* \tilde{W}^T(t) < 0$, from the bounds of the information matrix and the upper bound of the transition matrix as

$$\alpha I \leq \Omega(t) \leq \beta I, \qquad \Psi(t, t_j) \leq I, \qquad (48)$$

the following inequality is satisfied:

$$R_1(e, \tilde{W}) \leq f(e, \tilde{W}). \qquad (49)$$

It follows that if $F_1(e, \tilde{W})$ is positive semi-definite, the stability condition in Theorem 4 is satisfied, i.e. $F(e, \tilde{W}) \geq 0$. If $\Delta W_j^* \tilde{W}^T(t) \geq 0$, $F_1(e, \tilde{W})$ is always positive semi-definite and from Proposition 1, the Lyapunov stability is guaranteed. Therefore, $F_1(e, \tilde{W}) \geq 0$ is the sufficient condition for the Lyapunov stability of $(e, \tilde{W})$. $\qquad\square$

Proposition 3. *The tracking error e and the parameter estimation error $\tilde{W}$ are Lyapunov stable, only if $F_2(e, \tilde{W}) \geq 0$ for $t \in [t_i, t_{i+1}]$, where the function $F_2(e, \tilde{W})$ is defined as:*

$$
\begin{aligned}
F_2(e, \tilde{W}) &= \frac{1}{2} e^T(t) Q e(t) I \\
&+ \beta \tilde{W}(t) \tilde{W}^T(t) + \alpha \sum_{j=1}^{i} e^{-k(t - t_j)} \Delta W_j^* \tilde{W}^T(t)
\end{aligned}
\qquad (50)
$$

Proof. If $\Delta W_j^* \tilde{W}^T(t) < 0$, from the bounds of the information matrix and the lower bound of the transition matrix in Eqn. (22), the following equation is satisfied.

$$F(e, \tilde{W}) \leq F_2(e, \tilde{W}) \qquad (51)$$

If the stability condition in Theorem 4 is satisfied, the function $F_2(e, \tilde{W})$ is positive semi-definite. If $\Delta W_j^* \tilde{W}^T(t) \geq$

0, $F_2(e, \tilde{W})$ is always positive semi-definite and is satisfied regardless of the stability of $(e, \tilde{W})$. Therefore, $F_2(e, \tilde{W}) \geq 0$ is the necessary condition for the Lyapunov stability of $(e, \tilde{W})$. $\square$

While the exact values of the functions $F_1(e, \tilde{W})$ and $F_2(e, \tilde{W})$ are not obtainable, the effect of the parameter change ΔW_i^* and the forgetting factor k on them is interpreted in the following remark.

Remark 1.
- *The parameter change ΔW_i^* in the direction of making $\Delta W_j^* \tilde{W}(t)^T \geq 0$ does not effect on the stability condition.*
- *If the parameter change ΔW_i^* lies in the direction of $\Delta W_j^* \tilde{W}^T(t) < 0$, increase in the size of ΔW_i^* reduces both $F_1(e, \tilde{W})$ and $F_2(e, \tilde{W})$, narrowing down both the necessary and sufficient stability condition.*
- *When $\Delta W_j^* \tilde{W}^T(t) < 0$, large forgetting factor k increases $F_2(e, \tilde{W})$, enlarging the region for satisfying the necessary condtion.*
- *Large forgetting factor k results in smaller lower bound of the information matrix, reducing the convergence rate of parameter estimation.*

The value of the forgetting factor k needs to be chosen accordingly, considering the expected change of parameter, its effect on the stability conditions, and the desired convergence rate.

V. NUMERICAL SIMULATIONS

A. Simulation Setup

The wing rock roll dynamics, a common application example of adaptive control for its nonlinearity, is considered. The dynamics is modeled as in [18].

$$\dot{\phi} = p$$
$$\dot{p} = I_{xx}^{-1} \frac{1}{2} \rho U_\infty^2 Sb(C_l + C_{l_{\delta_a}} \delta_a) = \Delta(x) + L_{\delta_a} \delta_a, \tag{52}$$

where ϕ and p are the roll angle and its rate, and δ_a and $C_{l_{\delta_a}}$ are the aileron deflection and its nondimensional effectiveness.

Defining the state vector as $x = [\phi \quad p]^T$, the basis vector and the time-varying parameter are modelled as:
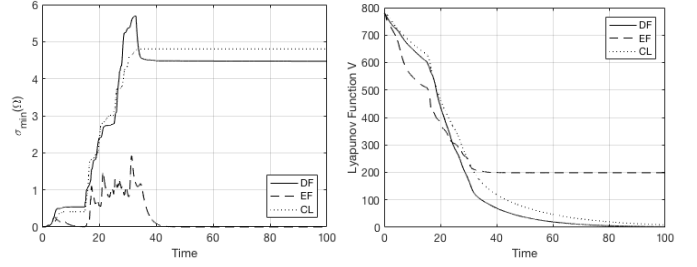
$$\begin{cases} \Phi(x) = [1, \phi, p, |\phi|p, |p|, \phi^3]^T, \quad t_1 = 50s \\ W_0^* = [.8, .2314, .6918, -.0624, .0095, .0215]^T \\ \Delta W_1^* = [.8, -.5 \cdot .2314, -.9 \cdot .6918, \\ \quad -9 \cdot .0624, 2 \cdot .0095, -.3 \cdot .0215]^T \end{cases} \tag{53}$$

The reference input is given as:

$$r(t) = g(t, 15) + g(t, 55) + g(t, 75) \\ + g(t, 95) + g(t, 115) + g(t, 135) \tag{54}$$

where the split square function $g(t, t_i)$ is defined as:

$$g(t, t_i) = \begin{cases} 4 & \text{for } t_i < t \leq t_i + 2 \\ 0 & \text{for } t_i + 2 < t \leq t_i + 10 \\ -4 & \text{for } t_i + 10 < t \leq t_i + 12 \end{cases} \tag{55}$$

(a) Minimum eigenvalue of $\Omega(t)$ (b) Lyapunov function $V(t)$

Fig. 1: Without parameter jump or excitation after 50s

For the reference, the performance of the proposed adaptive control with the directional forgetting (DF) is compared with that of the integral based concurrent learning (CL) algorithm without forgetting algorithm and the exponential forgetting (EF).

Without any forgetting, the CL method results in infinitely large information matrix, and thus requires an stack-manager algorithm. The stack-manager algorithm used in this paper is given as:

$$\Omega(t + \Delta t)$$
$$= \begin{cases} q(t)q^T(t), & \text{if } \text{rank}(\Omega(t)) < \text{rank}(\Omega(t) + q(t)q^T(t)) \\ & \text{or } ||\Phi(z(t)) - \Phi_{last}|| / ||\Phi(z(t))|| \geq tol \\ \Omega(t), & \text{otherwise} \end{cases}$$
$$M(t + \Delta t)$$
$$= \begin{cases} q(t)c^T(t), & \text{if } \text{rank}(\Omega(t)) < \text{rank}(\Omega(t) + q(t)q^T(t)) \\ & \text{or } ||\Phi(z(t)) - \Phi_{last}|| / ||\Phi(z(t))|| \geq tol \\ M(t), & \text{otherwise} \end{cases} \tag{56}$$

The EF method is a forgetting method in uniform directions. The discrete dynamics of the EF method is given as:

$$\Omega(t + \Delta t)$$
$$= \begin{cases} q(t)q^T(t), & \text{if } \text{rank}(\Omega(t)) < \text{rank}(\Omega(t) + q(t)q^T(t)) \\ \mu\Omega(t) + q(t)q^T(t), & \text{otherwise} \end{cases}$$
$$M(t + \Delta t)$$
$$= \begin{cases} q(t)c^T(t), & \text{if } \text{rank}(\Omega(t)) < \text{rank}(\Omega(t) + q(t)q^T(t)) \\ \mu M(t) + q(t)c^T(t), & \text{otherwise} \end{cases} \tag{57}$$

Likewise, the discrete dynamics of the DF method is

$$\Omega(t + \Delta t)$$
$$= \begin{cases} q(t)q^T(t), & \text{if } \text{rank}(\Omega(t)) < \text{rank}(\Omega(t) + q(t)q^T(t)) \\ \mu\dfrac{\Omega(t)q(t)q^T(t)}{q^T(t)\Omega(t)q(t)}\Omega(t) + q(t)q^T(t), & \text{otherwise} \end{cases}$$
$$M(t + \Delta t)$$
$$= \begin{cases} q(t)c^T(t), & \text{if } \text{rank}(\Omega(t)) < \text{rank}(\Omega(t) + q(t)q^T(t)) \\ \mu\dfrac{\Omega(t)q(t)q^T(t)}{q^T(t)\Omega(t)q(t)}M(t) + q(t)c^T(t), & \text{otherwise} \end{cases} \tag{58}$$

where the forgetting factor $0 < \mu \leq 1$ corresponds to e^{-k} of the continuous dynamics. Considering that μ is usually selected between $0.95 - 0.99$ and the forgetting of the EF is much faster than the DF, the value of μ is chosen as 0.99 in EF, and 0.95 in DF for their similar performance.

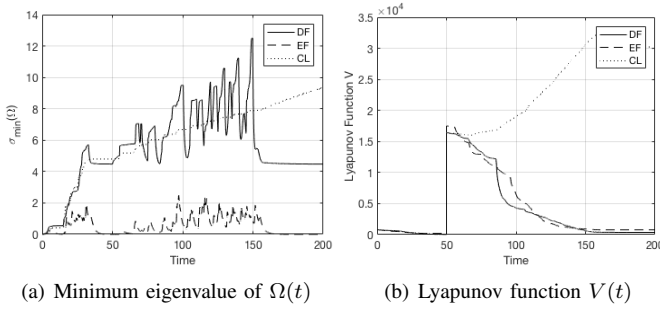(a) Minimum eigenvalue of $\Omega(t)$ (b) Lyapunov function $V(t)$

Fig. 2: With parameter jump at 50s

B. Simulation Results

First, assume there is no parameter change or reference inputs after 50s. The lower bound of the information matrix, i.e. the minimum singular value, and the Lyapunov function are shown in Fig. 1. The information matrix of DF is lower bounded, and both the tracking and parameter estimation error converge to zero without PE, while the information matrix of EF converges to zero resulting in the stagnation of the parameter estimation. The result clearly supports the Theorem 1 and 3.

If there is a parameter jump at 50s, the lower bound of the information matrix and the Lyapunov function are shown in Fig. 2. The information matrix of DF is lower bounded with a non-zero value as proven in Theorem 1, and upper bounded to a finite value as in Theorem 2. While the parameter error increases in CL, both forgetting algorithms show Lyapunov stability.

The result of the same simulation with different μ's is shown in Fig. 3, where the necessary condition for stability $F_2(e, \tilde{W})$ is given from Proposition 3. As the lower and upper bounds of the information matrix are not determined, worst-case values for reducing the $F_2(e, \tilde{W})$, i.e. $\alpha = \beta$, are chosen. The Lyapunov function is stable, and the necessary condition is met with positive semi-definiteness. Increase in k, or alternatively decrease in μ, results in large $F_2(e, \tilde{W})$ in the first part as the positive terms decay fast. In the later part, the parameter estimation is stagnated as the information matrix is lower bounded by a smaller value, which mainly determines the convergence rate

VI. CONCLUSION

In this paper, a directional forgetting based concurrent learning adaptive control has been proposed. The theoretical studies have shown that the information matrix is bounded, and both the tracking and parameter error converges to zero without the PE requirement under the assumption that there is no parameter change. The conditions for the convergence with the existence of parameter jumps have been obtained. The theoretical studies are supported by the numerical simulations on wingrock model. The proposed method is expected to be applicable to many adaptive control problems with time-varying parameters for its simplicity and convergence guarantee.

REFERENCES

[1] K. Astrom, B. Wittenmark, Adaptive control (2nd ed.), Reading, MA: Addsion-Wesley, 1995.
[2] G. Chowdhary, T. Yucelen, M. Muhlegg, E. N. Johnson, Concurrent Learning Adaptive Control of Linear Systems with Exponentially Convergent Bounds, International Journal of Adaptive Control and Signal Processing 22 (4) (2011) 325–343.
[3] G. Chowdhary, M. Mühlegg, E. Johnson, Exponential parameter and tracking error convergence guarantees for adaptive controllers without persistency of excitation, International Journal of Control 87 (8) (2014) 1583–1603.
[4] A. Parikh, R. Kamalapurkar, W. E. Dixon, Integral Concurrent Learning: Adaptive Control with Parameter Convergence without PE or State Derivatives, arXiv preprint arXiv:1512.03464.
[5] N. Cho, Y. Kim, Basis Integral Concurrent-Learning Model Reference Adaptive Control, 2016 European Control Conference.
[6] N. Cho, H.-S. Shin, Y. Kim, A. Tsourdos, Composite Model Reference Adaptive Control with Parameter Convergence under Finite Excitation, IEEE Transactions on Automatic Control.
[7] K. Oda, H. Takeuchi, M. Tsujii, M. Ohba, Practical Estimator for Self-Tuning Automotive Cruise Control, 1991 American Control Conference.
[8] N. Yoshitani, A. Hasegawa, Model-based control of strip temperature for the heating furnace in continuous annealing, IEEE Transactions on Control Systems Technology 6 (2) (1998) 146–156.
[9] A. Vahidi, A. Stefanopoulou, H. Peng, Recursive least squares with forgetting for online estimation of vehicle mass and road grade: theory and experiments, Vehicle System Dynamics 43 (1) (2005) 31–55.
[10] F. Fraccaroli, A. Peruffo, M. Zorzi, A new recursive least squares method with multiple forgetting schemes, Proceedings of the IEEE Conference on Decision and Control (2016) 3367–3372.
[11] R. M. Johnstone, C. Richard Johnson, R. R. Bitmead, B. D. Anderson, Exponential convergence of recursive least squares with exponential forgetting factor, Systems and Control Letters 2 (2) (1982) 77–82.
[12] Q. Zhang, A. Clavel, Adaptive observer with exponential forgetting factor for linear time varying systems, IEEE Conference on Decision and Control (CDC) 4 (December) (2001) 3886–3891.
[13] V. Adetola, M. Guay, Performance Improvement in Adaptive Control of Linearly Parameterized Nonlinear Systems, IEEE Transactions on Automatic Control 55 (9) (2010) 2182–2186.
[14] S. Bittanti, P. Bolzern, M. Campi, Convergence and exponential convergence of identification algorithms with directional forgetting factor, Automatica 26 (5) (1990) 929–932.
[15] L. Cao, H. Schwartz, Directional forgetting algorithm based on the decomposition of the information matrix, Mediterranean Conference on Control and Automation (MED09) (1999) 1635–1644.
[16] L. Cao, H. Schwartz, Directional forgetting algorithm based on the decomposition of the information matrix, Automatica 36 (11) (2000) 1725–1731.
[17] G. Tao, Adaptive control design and analysis, New York NY: Wiley, 2003.
[18] P. Ghorawat, Adaptive, Neural and Robust Control of Wing-Rock and Aeroelastic System, Ph.D. thesis, University of Nevada (2015).

Hae-In Lee received the BSc and MSc degrees in aerospace engineering from the Korea Advanced Institute of Science and Technology in 2013 and 2015, respectively. Since 2015, she has been a Ph.D. student in the School of Aerospace, Transport and Manufacturing at Cranfield University. Her primary research interests include networked control systems, adaptive control, and multi-objective optimisation.

Hyo-Sang Shin received his BSc from Pusan National University in 2004 and gained an MSc on flight dynamics, guidance and control in Aerospace Engineering from KAIST and a PhD on cooperative missile guidance from Cranfield University in 2006 and 2010, respectively. He is currently a Reader on Guidance, Control and Navigation Systems in Autonomous and Intelligent Systems Group at Cranfield University. His current research interests include nonlinear control, adaptive control, reconfigurable guidance and control, and cooperative and coordinated control of multiple vehicles.
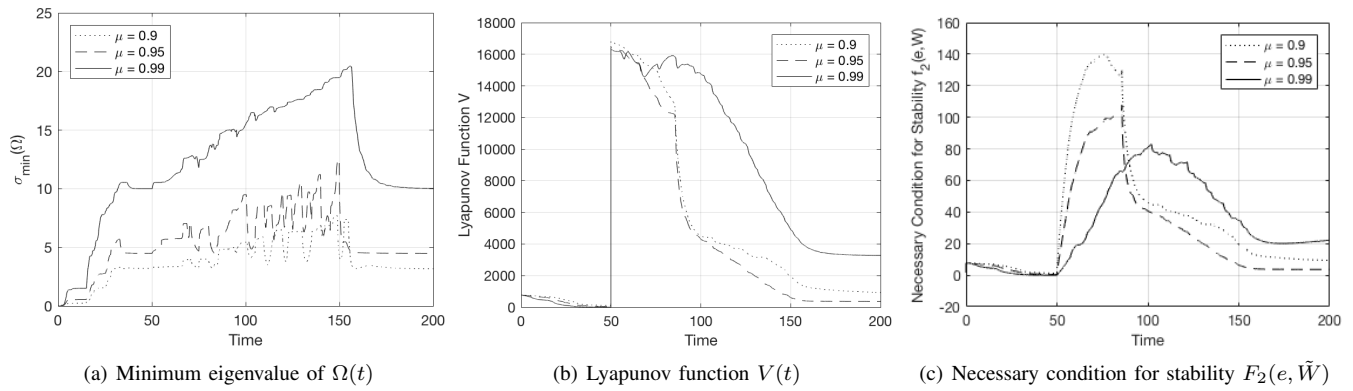
(a) Minimum eigenvalue of $\Omega(t)$ (b) Lyapunov function $V(t)$ (c) Necessary condition for stability $F_2(e,\tilde{W})$

Fig. 3: With parameter jump at 50s

Antonios Tsourdos obtained a MEng in electronic, control and systems engineering from the University of Sheffield (1995), an MSc in systems engineering from Cardiff University (1996), and a PhD in nonlinear robust missile autopilot design and analysis from Cranfield University (1999). He is a Professor of Control Engineering with Cranfield University, and was appointed Head of the Centre for Cyber-Physical Systems in 2013. He was a member of the Team Stellar, the winning team for the UK MoD Grand Challenge (2008) and the IET Innovation Award (Category Team, 2009).