IEEE Transactions on Automatic Control, Volume 68, Issue 5, May 2023, pp. 3086-3090
DOI:10.1109/TAC.2022.3197702


# Data-Efficient Active Weighting Algorithm for Composite Adaptive Control Systems

Seong-hun Kim, Hanna Lee, Namhoon Cho, *Member, IEEE*,
and Youdan Kim, *Senior Member, IEEE*

**Abstract**

We propose an active weighting algorithm for composite adaptive control to reduce the state and estimate errors while maintaining the estimation quality. Unlike previous studies that construct the composite term by simply stacking, removing, and pausing observed data, the proposed method efficiently utilizes the data by providing a theoretical set of weights for observations that can actively manipulate the composite term to have desired characteristics. As an example, a convex optimization formulation is provided, which maximizes the minimum eigenvalue while keeping other constraints, and an illustrative numerical simulation is also presented.

**Index Terms**

Composite adaptive control, Parameter estimation, Rank-one update.

## I. Introduction

Let us consider the following general form of composite model reference adaptive control (CMRAC) systems that can be found in [1]–[10].

$$\dot{e} = Ae + B\big(\delta(t,x) - W^T\phi(x)\big), \tag{1a}$$

$$\dot{W} = \Gamma\phi(x)e^T PB - (F(t)W - G(t)). \tag{1b}$$

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2019R1A2C208394612).

S-h. Kim, H. Lee, and Y. Kim are with the Institute of Advanced Aerospace Technology, Department of Aerospace Engineering, Seoul National University, Seoul, 08826, Republic of Korea (e-mail: ydkim@snu.ac.kr).

N. Cho is with Centre for Autonomous and Cyber-Physical Systems, School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, United Kingdom.

Here, $e \in \mathbb{R}^{n_e}$ is the tracking error vector, $x \in \mathcal{X} \subset \mathbb{R}^{n_x}$ is an exogenous state vector that may contain the states of a plant and/or a reference model, $W \in \mathbb{R}^{n \times m}$ denotes the adaptive parameter, $A \in \mathbb{R}^{n_e \times n_e}$ is Hurwitz, and $B \in \mathbb{R}^{n_e \times m}$ has full column rank. Let $\mathbb{S}^n$, $\mathbb{S}^n_+ \subset \mathbb{S}^n$ and $\mathbb{S}^n_{++} \subset \mathbb{S}^n$ denote a set of symmetric, positive semidefinite, and positive definite matrices in $\mathbb{R}^{n \times n}$, respectively. The matrix $\Gamma \in \mathbb{S}^n_{++}$ denotes the adaptation gain. The second term in the right of (1b) is the composite term composed of the functions $F$ and $G$, which are matrix-valued, piecewise continuous functions in $t \geq 0$, and $F(t) \in \mathbb{S}^n_+$. The matrix $P \in \mathbb{S}^{n_e}_{++}$ is the solution to the Lyapunov equation, $PA + A^T P + Q = 0$, where $Q \in \mathbb{S}^{n_e}_+$, and $(A, Q)$ is observable. Let $\mathbb{R}_+$ and $\mathbb{R}_{++}$ denote a set of nonnegative and positive real numbers, respectively. The function $\phi \colon \mathcal{X} \to \mathbb{R}^n$ is the known basis function locally Lipschitz in $x \in \mathcal{X}$, and $\delta \colon \mathbb{R}_+ \times \mathcal{X} \to \mathbb{R}^m$ denotes the time-varying, state-dependent uncertainty, which satisfies the following assumption.

**Assumption 1.** *The uncertainty function $\delta$ in* (1) *is linearly parameterized as*

$$\delta(t, x) = W^{\circ T} \phi(x) + \varepsilon(t, x), \tag{2}$$

*where $W^\circ \in \mathcal{W} \subset \mathbb{R}^{n \times m}$ is the unknown true constant parameter, and $\varepsilon : \mathbb{R}_+ \times \mathcal{X} \to \mathbb{R}^m$ is the parameterization residue which is piecewise continuous in $t$ and locally Lipschitz in $x$, and bounded by a positive scalar as $\|\varepsilon(t, x)\| \leq b_\varepsilon$ for all $t \in \mathbb{R}_+$ and $x \in \mathcal{X}$.*

The composite term $F(t)W - G(t)$ is typically designed to represent the gradient of a convex loss function $l \colon \mathcal{W} \to \mathbb{R}$ which hopefully has one of its minimum at the true parameter $W^\circ$ in (2). Since it is difficult to obtain the uncertainty $\delta(t, x)$ in (2) directly, an observer or a differentiator is typically used to estimate $W^\circ$ by observing $(\xi, y)$ satisfying

$$y(t) = W^{\circ T} \xi(t) + \epsilon(t), \tag{3}$$

where $\epsilon : \mathbb{R}_+ \to \mathbb{R}^m$ denotes the estimation residue bounded by $\|\epsilon(t)\| \leq b_\epsilon$ for all $t \geq 0$. This estimation residue may exist even if the parameterization residue $\varepsilon \equiv 0$ in (2).

Without the composite term, the dynamics (1) becomes vanilla MRAC where it guarantees the uniform ultimate boundedness (UUB) [11] of $\tilde{W} \coloneqq W - W^\circ$, and $e \to 0$ by Barbalat's Lemma [12]–[14]. The CMRAC utilizes the composite term to improve the vibrating responses in the vanilla MRAC, by converging not only $\|e\|$ but also $l$ to their minimum [1]–[5]. The above methods require the persistently excited (PE) $\phi$ for exponential convergence of $e$ and $\tilde{W}$, and the UUB of them under $\varepsilon \neq 0$ [1], [12], [13]. Since the PE condition is difficult to guarantee and calculate, memory-based CMRACs utilizing historical $(\xi, y)$ stored in the memory have been developed over the past decade [6]–[10]. If the stored data is used in batches to obtain the strong convexity of $l$, or equivalently $F(t) \in \mathbb{S}^n_{++}$, then $e$ and $\tilde{W}$ have the UUB and exponential convergence. For example, the concurrent learning methods in [6], [7] use $F(t) = \sum_{i=1}^{N} \xi(t_i)\xi(t_i)^T$ and $G(t) = \sum_{i=1}^{N} \xi(t_i)y(t_i)^T$, where $N$ denotes the size of the memory. Integral-based methods in [8]–[10] use $F(t) = \int_0^t \gamma(t,s)\xi(s)\xi(s)^T \,\mathrm{d}s$ and $G(t) = \int_0^t \gamma(t,s)\xi(s)y(s)^T \,\mathrm{d}s$, where $\gamma$ denotes the forgetting factor. Let $\lambda_1$ be the minimum eigenvalue of $F(t)$. Memory-based CMRACs are equipped with algorithms that can monotonically increase $\lambda_1$ to ensure $F(t) \in \mathbb{S}^n_{++}$ after some time instance. Moreover, since the larger the $\lambda_1$, the smaller the UUB, the main concern of the algorithms is how to increase $\lambda_1$ as much and/or as quickly as possible. However, most of the existing algorithms rely on passive methods waiting for new $(\xi, y)$ that can increase $\lambda_1$ by chance, which are inefficient ways of utilizing data.

In this technical note, we propose an algorithm that can efficiently manipulate the characteristics of $F(t)$ and $G(t)$. First, we show that not only $\lambda_{\min}(F(t))$ but also $G(t) - F(t)W$ determine the UUB when $\varepsilon \neq 0$. The purpose of the algorithm is to minimize the UUB and to bound the condition number of $F(t)$, that determines the accuracy of estimate $W$ when $\varepsilon \neq 0$ [15], through actively weighting $(\xi, y)$ in the memory. The rank-one update theory in [16] guarantees that there exists a convex set of weights satisfying all the desired conditions of $F(t)$ and $G(t)$. The weights can then be chosen from the set to conform to the desired strategy, such as to maximize the minimum eigenvalue while keeping the maximum eigenvalue and the condition number. We provide the convex optimization formulation to implement this strategy, as an example.

This technical note is organized as follows. The basic notations used in this study are summarized in Section II. The memory-base CMRAC and its UUB and the condition number analysis are given in Section III. Section IV introduces the active weighting algorithm, and the corresponding excitation condition, and Section V presents the convex optimization problem. An illustrative numerical example is provided in Section VI, and the conclusions are summarized in Section VII.

## II. Preliminaries

Let $I_n \in \mathbb{R}^{n \times n}$ denote the identity matrix, $A \succeq 0$ denote $A \in \mathbb{S}^n$, and a relation $A \succeq B$ for $A, B \in \mathbb{S}^n$ mean that $A - B \succeq 0$. The norm of a matrix $A$ is defined by the Frobenius norm as $\|A\| = \sqrt{\mathrm{tr}(A^T A)}$, where $\mathrm{tr}(\cdot)$ denotes the trace operator, and the norm of a vector

$x$ is defined by the Euclidean norm as $\|x\| = \|x\|_2$. For two normed vector spaces $X$ and $Y$, define the norm on the Cartesian product $X \times Y$ as $\|(x,y)\| = \sqrt{\|x\|^2 + \|y\|^2}$, where $(x,y) \in X \times Y$. Given $F \in \mathbb{S}^n$, let $\lambda_i(F)$ denote the $i$-th eigenvalue of $F$ in increasing order, i.e., $\lambda_{\min}(F) = \lambda_1(F) \leq \ldots \leq \lambda_n(F) = \lambda_{\max}(F)$. Let $v_i$'s denote the orthonormal eigenvectors of $F$. Then, the projection of a vector $\xi \in \mathbb{R}^n$ is defined by

$$\mathcal{P}_{i:j}^F(\xi) = \left[v_i \cdots v_j\right]^T \xi, \quad \forall 1 \leq i \leq j \leq n. \tag{4}$$

To simplify the notation, let $\xi_{i:j}$ stand for $\mathcal{P}_{i:j}^F(\xi)$, and let $\xi_i = \xi_{i:i}$. If there exists a degenerated eigenvalue $\lambda_i(F)$ that has the multiplicity, denoted by $\rho_i(F)$, greater than one, then we list the eigenvalues in the order in which $|\xi_i|$'s do not decrease as follows.

$$\lambda_{i-1}(F) < \lambda_i(F) = \ldots = \lambda_{i+\rho_i(F)-1}(F) < \lambda_{i+\rho_i(F)}, \tag{5}$$

where $|\xi_i| \leq \ldots \leq \left|\xi_{i+\rho_i(F)-1}\right|$. The condition number of $G \in \mathbb{S}_{++}^n$ is denoted by $\kappa(G) \coloneqq \lambda_n(G)/\lambda_1(G)$. The binary operators $\wedge$ and $\vee$ denote the minimum and maximum operators, respectively, such that $a \wedge b = \min(a,b)$ and $a \vee b = \max(a,b)$ for $a, b \in \mathbb{R}$.

## III. Memory-Based CMRAC

This section introduces a memory-based composite adaptation algorithm that constructs $F$ and $G$ in (1b). Also, the Lyapunov stability analysis is given for the systems in (1). Suppose that the controller observes data at each time $t \in \mathcal{T} \coloneqq \{t_i\}$, where $i \in \mathbb{N}$ and $t_1 = 0$. The data at time $t$ is denoted by $(\xi(t), y(t))$, which satisfies the relation in (3). For each time $t \in \mathbb{R}_+$, the only data that the controller is supplied with is determined by an index set; this set is denoted as $\mathcal{I}_k \coloneqq \{i\}_{i=1}^k$ for $k \in \mathbb{N}$, such that $t \in [t_k, t_{k+1})$, where $t_k, t_{k+1} \in \mathcal{T}$. Consider the following two auxiliary matrices.

$$F_k \coloneqq \sum_{i \in \mathcal{I}_k} \gamma_i^{(k)} \xi(t_i) \xi(t_i)^T, \tag{6a}$$

$$G_k \coloneqq \sum_{i \in \mathcal{I}_k} \gamma_i^{(k)} \xi(t_i) y(t_i)^T, \tag{6b}$$

where the scalar $\gamma_i^{(k)}$ denotes the weight given to the data measured at time $t_i \leq t_k$ in the time interval $[t_k, t_{k+1})$.

**Theorem 1.** *For all $k \in \mathbb{N}$, suppose that $F$ and $G$ in* (1) *satisfy $F(t) = F_k$ if $\lambda_1(F_k) \geq b_f$, $F(t) = b_f I_n$ otherwise, and $G(t) = G_k$ for all $t \in [t_k, t_{k+1})$ with $b_f > 0$ and* (6)*. Under Assumption 1, if the weights $\gamma_i^{(k)}$ satisfy*

$$b_F I_n \succeq F_{k+1} \succeq \lambda_1(F_k) I_n \succeq 0, \tag{7a}$$

$$h_k := \sum_{i \in \mathcal{I}_k} \left|\gamma_i^{(k)}\right| \|\xi(t_i)\| \leq b_\gamma, \tag{7b}$$

*where $b_F, b_\gamma \in \mathbb{R}_{++}$, then, the condition number $\kappa(F(t)) \leq b_F / b_f$ for all $t \geq 0$, and the error $\eta(t) := (e(t), \tilde{W}(t))$ of* (1) *is globally uniformly ultimately bounded. Moreover, if there is $T_1 \geq 0$ such that $M := \lambda_1(F(T_1)) > b_f$, then the condition number $\kappa(F(t)) \leq b_F / M$ for all $t \geq T_1$, and the error $\eta(t)$ satisfies that*

$$\|\eta(t)\| \leq \frac{b}{\theta}\left(\frac{2}{\lambda_1(Q)} + \frac{\lambda_n(\Gamma)}{M}\right)\sqrt{\frac{c_2}{c_1}}, \tag{8}$$

*for all $t \geq T_2$ with a finite time $T_2 \geq T_1$, where $\theta \in (0,1)$, $c_1 := \lambda_1(P) \wedge \lambda_1(\Gamma^{-1})$, $c_2 := \lambda_{n_e}(P) \vee \lambda_n(\Gamma^{-1})$, and*

$$b := \sqrt{b_E^2 \lambda_n(\Gamma^{-1})^2 + b_\varepsilon^2 \|PB\|^2}, \tag{9}$$

*where $b_E := b_\gamma b_\epsilon + b_f \|W^\circ\|$.*

*Proof.* From (7a), $\lambda_1(F(t))$ is monotonically increasing with $t$, and it is bounded from below by $b_f$. It follows that $\kappa(F(t)) \leq b_F b_f^{-1}$ for all $t \geq 0$ and $\kappa(F(t)) \leq b_F M^{-1}$ for all $t \geq T_1$ from (7a). Since $\left\|y(t_i) - W^{\circ T}\xi(t_i)\right\| = \|\epsilon(t_i)\| \leq b_\epsilon$, that $\|G(t) - F(t)W^\circ\| \leq b_E$ follows from (7b). Let a Lyapunov function candidate be

$$V(t, \eta) = \frac{1}{2} e^T P e + \frac{1}{2} \operatorname{tr}\left(\tilde{W}^T \Gamma^{-1} \tilde{W}\right). \tag{10}$$

By [11, Section 4.8], the time-derivative of $V$ satisfies $\dot{V}(t,\eta) \leq -(1-\theta)\alpha(b_f)\|\eta\|^2$, for $\|\eta\| \geq \frac{b}{\theta\alpha(b_f)}$, where the function $\alpha : \mathbb{R}_+ \to \big[0, \lambda_1(Q)/2\big)$ is defined by

$$\alpha(b_f) = \frac{b_f \lambda_1(\Gamma^{-1})\lambda_1(Q)}{2 b_f \lambda_1(\Gamma^{-1}) + \lambda_1(Q)}. \tag{11}$$

By [11, Theorem 4.18], there exists a finite $t_1 \geq t_0$ for any $\eta(t_0)$ such that the solution $\eta(t)$ for all $t \geq t_1$ satisfies $\|\eta(t)\| \leq \frac{b}{\theta\alpha(b_f)}\sqrt{\frac{c_2}{c_1}}$, which implies the global UUB of $\eta$.

If there exist $T_1$ and $M$ given in the condition, we have $\alpha(M) > \alpha(b_f)$, and by using the similar approach used above, there exists a finite $t_2 \geq 0$ such that for all $t \geq T_2 := t_2 + T_1 \geq T_1$, $\|\eta(t)\| \leq \frac{b}{\theta\alpha(M)}\sqrt{\frac{c_2}{c_1}}$, which completes the proof. □

According to Theorem 1, the scalars $b_f$, $b_F$, $b_\gamma$, and $\lambda_1(F_k)$ determine $\kappa(F(t))$ and the UUB of $\eta(t)$. The scalar $b_f$ is an auxiliary parameter that limits the lower bound of $\lambda_1(F(t))$ to a positive number for all $t \geq 0$; this is similar to the $\sigma$-modification MRAC used when the signal is not PE [7]. The condition $\lambda_1(F_k) \geq 0$ has been considered in many studies [7]–[10]. However, the scalars $b_F$ and $b_\gamma$ have not been addressed much in the literature. The condition number of $F(t)$ may affect the estimation residue $\epsilon(t)$ in (3) when $\varepsilon \neq 0$ [15]. Hence, for a small UUB in (8), not only $b_\gamma$ but also $b_F$ should be kept as small as possible to make the condition number $\kappa(F(t))$, and therefore $b_\epsilon$ small.

## IV. Active Weighting Algorithm

In this section, the active weighting algorithm that keeps $b_F$ and $b_\gamma$ bounded while increasing $\lambda_1(F_k)$ as much as possible is proposed. Let $\gamma_i^{(k)}$ be defined by $\gamma_i^{(k)} = p_i \prod_{j=i+1}^{k} q_j$ if $i < k$, and $\gamma_i^{(k)} = p_k$ if $i = k$, for all $i \in \mathcal{I}_k$ and $k \in \mathbb{N}$ with some sequences of real numbers $\{p_i\}$ and $\{q_i\}$. This form of the weights allows us to draw the following simplified recursive rank-one update laws from (6) and (7b) as

$$F_+ = qF + p\xi\xi^T, \quad h_+ = qh + |p|\|\xi\|, \tag{12}$$

where $p \in \mathbb{R}$ and $q > 0$ denote the weight of update (12), $F \in \mathbb{S}_+^n$, $\xi \in \mathbb{R}^n$, $\lambda_n(F) \leq b_F$, $0 \leq h \leq b_h$, and the corresponding update law for $G$ is given by

$$G_+ = qG + p\xi y^T. \tag{13}$$

### A. Set of Weights

Given (12), the aim of this section is to find a set of $(p, q)$ satisfying

$$\mu_F \leq \lambda_i(F_+) \leq L_F, \quad 0 \leq h_+ \leq L_h, \tag{14}$$

for all $1 \leq i \leq n$ and some nonnegative scalars $\mu_F$, $L_F$, and $L_h$. To this end, the theoretical findings in [16] is summarized in the following lemma with a slight modification due to $p$ and

$q$. Given $F \in \mathbb{S}_+^n$ and $\xi \in \mathbb{R}^n$, let us define the corresponding functions $f_i^\pm$ and $g_j^\pm$ as

$$\begin{aligned} f_i^\pm(p,q) &:= \frac{1}{2}\big(\|\xi_{i:n}\|^2 p + \bar{\lambda}_i(F) q\big) \\ &\quad \pm \frac{1}{2}\left\| \begin{bmatrix} \|\xi_{i+1:n}\|^2 - |\xi_i|^2 & \tilde{\lambda}_i(F) \\ 2|\xi_i|\|\xi_{i+1:n}\| & 0 \end{bmatrix} \begin{pmatrix} p \\ q \end{pmatrix} \right\|, \end{aligned} \tag{15a}$$

$$\begin{aligned} g_j^\pm(p,q) &:= \frac{1}{2}\big(\|\xi_{1:j}\|^2 p + \bar{\lambda}_{j-1}(F) q\big) \\ &\quad \pm \frac{1}{2}\left\| \begin{bmatrix} |\xi_j|^2 - \|\xi_{1:j-1}\|^2 & \tilde{\lambda}_{j-1}(F) \\ 2|\xi_j|\|\xi_{1:j-1}\| & 0 \end{bmatrix} \begin{pmatrix} p \\ q \end{pmatrix} \right\|, \end{aligned} \tag{15b}$$

for all $i, j \in \mathbb{N}$ such that $1 \le i \le n-1$, $2 \le j \le n$, $\bar{\lambda}_i(F) := \lambda_{i+1}(F) + \lambda_i(F)$, and $\tilde{\lambda}_i(F) := \lambda_{i+1}(F) - \lambda_i(F)$.

**Lemma 1** (Ipsen and Nadler [16])**.** *Given $F \in \mathbb{S}_+^n$ and $\xi \in \mathbb{R}^n$, the updated matrix $F_+$ in* (12) *satisfies the following conditions for any $p \in \mathbb{R}$ and $q > 0$, with $f_i^\pm$ and $g_j^\pm$ defined in* (15a) *and* (15b)*, respectively.*

(i) $f_1^-(p,q) \le \lambda_1(F_+) \le g_2^-(p,q)$,

(ii) $f_{n-1}^+(p,q) \le \lambda_n(F_+) \le g_n^+(p,q)$,

(iii) $f_i^-(p,q) \vee q\lambda_{i-1}(F) \le \lambda_i(F_+) \le g_i^+(p,q) \wedge q\lambda_{i+1}(F)$, *for* $2 \le i \le n-1$.

Let $\Omega_i(\mu_F, L_F, L_h)$ be a set of $(p,q) \in \mathbb{R} \times \mathbb{R}_{++}$ such that $f_i^-(p,q) \ge \mu_F$, $g_n^+(p,q) \le L_F$, and $|p|\|\xi\| + qh \le L_h$. The set $\Omega_i(\mu_F, L_F, L_h)$ is convex for all $i$. Moreover, if there exists $(p,q) \in \Omega_1(\mu_F, L_F, L_h)$, then (14) is satisfied under (12) by Lemma 1. If there exists $(p_k, q_k) \in \Omega_1(\lambda_1(F_k), b_F, b_h)$ for all $k$, the iteration (12) implies the hypotheses of Theorem 1.

**Proposition 1.** *Let $x_j, y_j \in \mathbb{R}_+$ for $j = 1,2,3$. Then, for all $1 \le i \le n$, $\Omega_i(x_1,x_2,x_3) \subset \Omega_i(y_1,y_2,y_3)$, if and only if $x_1 \ge y_1$, $x_2 \le y_2$, and $x_3 \le y_3$.*

In addition to satisfying (14), another goal of the algorithm is to increase $\lambda_1(F_+)$ as fast as possible so that there is no eigenvalue of zero. However, it will be shown that if $\rho_1(F) \ge 2$, only one of them can be increased with (12). Let us introduce a practical assumption before continuing the analysis.

**Assumption 2.** *All nonzero eigenvalues of $F$ in* (12) *are distinguishable.*

The multiplicity of eigenvalue zero is inevitable in the early stages of the iteration because the update (12) starts with $F = 0$. However, the distinct nonzero eigenvalues are generally acceptable since $F$ is updated by $\xi$ that depends on the state $x$ of a real physical plant. Moreover, adjusting $p$ and $q$ can also avoid the situation that $F$ has indistinguishable nonzero eigenvalues.

**Definition 1** (Minimum Effective Eigenvalue)**.** *Given $F \in \mathbb{S}_+^n$ and $\xi \in \mathbb{R}^n$, let $r = \rho_1(F)$. Then, the eigenvalue $\lambda_r(F)$ is the minimum effective eigenvalue of $F$ under $\xi$.*

When $\rho_1(F) = 1$, the minimum effective eigenvalue is always the unique minimum eigenvalue which can be zero or not. However, $\rho_1(F) \geq 2$ implies that there exist two or more eigenvalues of zero, and the one with the largest $|\xi_i|$ is the minimum effective eigenvalue.

**Theorem 2.** *Given* (12) *under Assumption 2, let $r = \rho_1(F)$. Then, for all $(p, q) \in \Omega_r(\mu_F, L_F, L_h)$, it follows that $\lambda_r(F_+) \geq \mu_F$, and $\lambda_i(F_+) = 0$ for all $1 \leq i \leq r-1$.*

*Proof.* First, we will show that $\lambda_r(F_+) \geq \mu_F$. When $r = 1$, it follows immediately that $\lambda_r(F_+) = \lambda_1(F_+) \geq f_1^-(p, q) \geq \mu_F$ by Lemma 1 (i). When $r \geq 2$, we have $\lambda_r(F_+) \geq f_r^-(p, q) \vee 0 \geq f_r^-(p, q) \geq \mu_F$ by Lemma 1 (iii).

Now, let us show that $\lambda_i(F_+) = 0$ for all $1 \leq i \leq r-1$. By Lemma 1 (i), we have

$$0 \leq f_1^-(p, q) \leq \lambda_1(F_+) \leq g_2^-(p, q) = \frac{1}{2}\|\xi_{1:2}\|^2(p - |p|) \leq 0, \tag{16}$$

which implies that $\lambda_1(F_+) = 0$. For $2 \leq i \leq r-1$, we have $\lambda_{i-1}(F) = \lambda_{i+1}(F) = 0$, and the corresponding functions satisfy $f_i^-(p, q) = \frac{1}{2}\|\xi_{i:n}\|^2(p - |p|) \leq 0$ and $g_i^+(p, q) = \frac{1}{2}\|\xi_{1:i}\|^2(p + |p|) \geq 0$ for all $p \in \mathbb{R}$. Hence, from Lemma 1 (iii), we have

$$0 = f_i^-(p, q) \vee q\lambda_{i-1}(F) \leq \lambda_i(F_+) \leq g_i^+(p, q) \wedge q\lambda_{i+1}(F) = 0, \tag{17}$$

which completes the proof. □

Theorem 2 indicates that only the minimum effective eigenvalue can be increased in one update in (12) for any $\mu_F$, $L_F$, and $L_h$, because $\Omega_r(\mu_F, L_F, L_h) \subset \Omega_r(0, \infty, \infty)$ by Proposition 1.

**Remark 1.** *Even if Assumption 2 does not hold, i.e., there is a nonzero eigenvalue with multiplicity greater than* 1*, Lemma 1 still holds. In the worst-case scenario where the minimum eigenvalue is nonzero and has the multiplicity greater than one, only one of them can increase while the others are unchanged, as in the case of the eigenvalue of zero in Theorem 2, with an additional constraint $q \geq 1$. This constraint reduces the set of weight $(p, q)$, but may decrease* *the multiplicity. Therefore, Assumption 2 holds within a finite number of updates with some additional conditions on the observed data $(\xi, y)$. To avoid the complexity of considering all the situations and based on the practical reasons discussed in Section IV, Assumption 2 is considered throughout this study.*

### B. Excitation Condition

This section provides the condition that $\Omega_r(\mu_F, L_F, L_h)$ in Theorem 2 is nonempty. Indeed, if $L_F \geq \lambda_n(F)$ and $L_h \geq h$, the set $\Omega_r(\lambda_r(F), L_F, L_h)$ is nonempty because $(0, 1)$ is the trivial point of $\Omega_r(\lambda_r(F), \lambda_n(F), h) \subset \Omega_r(\lambda_r(F), L_F, L_h)$. However, for $(p, q) \in \Omega_r(\lambda_r(F), L_F, L_h)$, there may be the case that $\lambda_r(F_+) = \lambda_r(F)$ by Theorem 2, which is said that $\xi$ does not excite $F$. Therefore, we introduce an excitation condition for $\xi$ and $F$, which can be interpreted as the existence of nonempty set $\Omega_r(\mu_F, L_F, L_h)$ where $\mu_F > \lambda_r(F)$.

**Theorem 3.** *Given* (12) *under Assumption 2, suppose that $\mu_F$, $L_F$, and $L_h$ satisfy $\mu_F > \lambda_r(F)$, $L_F > \lambda_n(F)$, and $L_h > h$. Then, for all $\xi \in \mathbb{R}^n$ with $\xi_r \neq 0$ where $r = \rho_1(F)$, the set $\Omega_r(\mu_F, L_F, L_h)$ is nonempty.*

*Proof.* Let

$$p_0 = \frac{L_h - h}{2\|\xi\|} \wedge \frac{L_F - \lambda_n(F)}{2\|\xi\|^2}, \tag{18}$$

$$q_0 = \frac{L_h - p_0\|\xi\|}{h} \wedge \frac{L_F - p_0\|\xi\|^2}{\lambda_n(F)}. \tag{19}$$

Note that $(p_0, q_0) \in \Omega_r(\lambda_r(F), L_F, L_h)$, $p_0 > 0$, and

$$q_0 \geq \frac{L_h + h}{2h} \wedge \frac{L_F + \lambda_n(F)}{2\lambda_n(F)} > 1. \tag{20}$$

Hence, it follows that

$$\begin{aligned} f_r^-(p_0, q_0) &= \frac{1}{2}\big(\|\xi_{r:n}\|^2 p_0 + \bar{\lambda}_r(F) q_0\big) \\ &\quad - \frac{1}{2}\sqrt{\Big(\|\xi_{r:n}\|^2 p_0 + \tilde{\lambda}_r(F) q_0\Big)^2 - 4\tilde{\lambda}_r(F)|\xi_r|^2 p_0 q_0} \\ &> \frac{1}{2}\Big(\bar{\lambda}_r(F) - \tilde{\lambda}_r(F)\Big) q_0 = \lambda_r(F) q_0 > \lambda_r(F), \end{aligned} \tag{21}$$

which implies that there exists $\mu_F$ such that $f_r^-(p_0, q_0) \geq \mu_F > \lambda_r(F)$. With $\Omega_r(\mu_F, L_F, L_h) \subset \Omega_r(\lambda_r(F), L_F, L_h)$ by Proposition 1, we have $(p_0, q_0) \in \Omega_r(\mu_F, L_F, L_h)$. □

Since $L_F$ and $L_h$ are design parameters, it is always possible to find $(p, q) \in \Omega_r(\mu_F, L_F, L_h)$ where $\mu_F > \lambda_r(F)$ whenever $\xi_r \neq 0$ by Theorem 3, and therefore the update (12) implies $\lambda_r(F_+) > \lambda_r(F)$ by Theorem 2.

## V. Weight Selection Examples

Once $\mu_F$, $L_F$, and $L_h$ are determined to satisfy the hypothesis of Theorem 1, various algorithms, such as greedy or optimization methods, can be used to choose the weight $(p, q) \in \Omega_r(\mu_F, L_F, L_h)$ according to the strategy to be used. This section introduces a convex optimization formulation and its reduced form to find the optimal weight according to the strategy that maximizes $\lambda_r(F_+)$ while satisfying (7).

Consider the following optimization problem.

$$\begin{aligned} \underset{p,q,\mu_F}{\text{maximize}} \quad & \mu_F \\ \text{subject to} \quad & (p, q) \in \Omega_r(\mu_F, L_F, L_h), \end{aligned} \tag{22}$$

where $0 \leq L_F \leq b_F$ and $0 \leq L_h \leq b_h$. This is a SOCP because $(p, q) \in \Omega_r(\mu_F, L_F, L_h)$ is a second-order cone constraint by (14) and (15), and the solution satisfying $\mu_F > \lambda_r(F)$ can always be found under the hypotheses of Theorem 3.

The optimization problem (22) can further be reduced to a simpler line search problem. From (15b), if we fix $p$, there is a unique $q > 0$ such that $g_n^+(p, q) = L_F$. Let $q_F(p)$ be the unique solution, and

$$q_h(p) := \frac{L_h - |p| \|\xi\|}{h}. \tag{23}$$

Also, we have $p \in (-L_h/\|\xi\|, L_h/\|\xi\|)$ from (14).

**Lemma 2.** *For the problem described by* (22), *let* $\Omega = \Omega_r(\lambda_r(F), L_F, L_h)$, *and* $\xi_r \neq 0$. *Then,* $(p^*, q^*) := \arg\max_{(p,q)\in\Omega} f_r^-(p, q)$ *satisfies* $q^* = q_F(p^*) \wedge q_h(p^*)$.

*Proof.* Let a matrix $C$ be

$$C = \begin{bmatrix} \|\xi_{r+1:n}\|^2 - |\xi_r|^2 & \tilde{\lambda}_r(F) \\ 2|\xi_r|\|\xi_{r+1:n}\| & 0 \end{bmatrix}. \tag{24}$$

From (15a), the gradient of $f_r^-$ with respect to $\zeta := [p, q]^T$ is given by

$$\boldsymbol{\nabla} f_r^-(p, q) = \frac{1}{2} \begin{pmatrix} \|\xi_{r:n}\| \\ \bar{\lambda}_r(F) \end{pmatrix} - \frac{1}{2} \|C\zeta\|^{-1} C^T C \zeta, \tag{25}$$

and this implies that $f_r^-(p,q) = \zeta^T \nabla f_r^-(p,q)$ for all $(p,q) \in \mathbb{R} \times \mathbb{R}_{++}$. It is easily seen that $f_r^-(p_0^*, q_0^*) = 0$, for all optimal points $(p_0^*, q_0^*)$ of $f_r^-$ such that $\nabla f_r^-(p_0^*, q_0^*) = 0$.

For $\lambda_r(F) > 0$, which means $r = 1$ from Assumption 2, suppose that $(p_0^*, q_0^*) \in \Omega$. From the definition, any tuple $(p,q) \in \Omega$ satisfies $f_r^-(p,q) \geq \lambda_r(F) > 0$. Thus, by contradiction, the tuple $(p_0^*, q_0^*) \notin \Omega$, which implies $(p^*, q^*)$ lies on the boundary of $\Omega$. Moreover, from Theorem 3, there always exists a tuple $(p', q')$ such that $f_r^-(p', q') > \lambda_r(F)$, thereby, it can be concluded that $(p^*, q^*)$ satisfies $q^* = q_F(p^*) \wedge q_h(p^*)$.

When $\lambda_r(F) = 0$, all the tuples $(p_0^*, q_0^*)$ are contained in a set $\{(p,q) \mid f_r^-(p,q) = \lambda_r(F) = 0\}$. Again, by Theorem 3, there always exists a tuple $(p', q')$ such that $f_r^-(p', q') > 0$. It follows that $(p^*, q^*) \notin \{(p,q) \mid f_r^-(p,q) = \lambda_r(F) = 0\}$. Therefore, we have $q^* = q_F(p^*) \wedge q_h(p^*)$, which completes the proof. $\square$

By Lemma 2, the SOCP in (22) can be reduced to the following line search problem which is much faster to solve.

$$\begin{aligned} &\underset{p}{\text{maximize}} && f_r^-(p, q_h(p) \wedge q_F(p)) \\ &\text{subject to} && p \in \left( -\frac{L_h}{\|\xi\|}, \frac{L_h}{\|\xi\|} \right). \end{aligned} \tag{26}$$

The singular value maximizing algorithm for concurrent learning methods executes the singular value decomposition $N$ times for each time $t \in \mathcal{T}$, where $N$ denotes the size of the memory [6], [7]. Since $N \geq n$ is required for $\lambda_r(F_k) > 0$ for some $k$, the time complexity of the algorithm is $\mathcal{O}(n^4)$ if the decomposition takes $\mathcal{O}(n^3)$ time. The proposed method, as well as the integral-based methods in [8]–[10], takes $\mathcal{O}(n^3)$ time, because it requires only one execution of the eigenvalue decomposition for each $t \in \mathcal{T}$ and the time complexity of finding $(p,q) \in \Omega_r(\mu_F, L_F, L_h)$ is linear in $n$. Note that $\mathcal{O}(n^3 + n) = \mathcal{O}(n^3)$. For example, Brent's method for solving the line search problem in (26) takes $\mathcal{O}(L^2 n)$ where $L$ is the maximum number of iteration steps of the optimization process, which is independent of $n$. For the same $n$, the integral-based methods such as [8]–[10] may take shorter time at each $t \in \mathcal{T}$ than the proposed method using optimization process. However, they still have the same linear time complexity in the number of observation steps which depends on the measurement update interval.

## VI. Numerical Example

To illustrate the characteristics of the composite term satisfying (7) under the hypotheses of Theorem 2, three CMRAC methods are compared using the same system as used in [10]: i) a concurrent learning-based CMRAC with singular value maximization [6], denoted by CL-CMRAC, ii) an integral-based CMRAC with finite excitation [10], denoted by FE-CMRAC, and iii) the proposed method. The data $(\xi, y)$ is assumed to be observed every $0.01$ seconds, and to illustrate the effectiveness of bounding the condition number $\kappa(F_k)$ and $h_k$, the following parameterization residue in (2) is considered.

$$\varepsilon(t,x) = \sin(5t) + \cos(2t)^2. \tag{27}$$

The bounds are set as $b_F = 2.637$ and $b_h = 10$, and $L_F$ and $L_h$ are gradually increased to these bounds for the proposed method to satisfy the hypotheses of Theorem 3. Brent's method is used to solve the line search problem (26). Also, $N = 10$ for CL-CMRAC, and all the other design parameters of CL-CMRAC and FE-CMRAC are set to be the same as in [10], except for the weights of each observation $\gamma_i^{(k)}$ in (6), which will be discussed later.

The state, control input, and the tracking error norm $\|\eta(t)\|$ responses are presented in Fig. 1, and the characteristics of composite term represented by $\lambda_1(F_k)$, $\kappa(F_k)$, and $h_k$ are shown in Fig. 2. The weights $\gamma_i^{(k)}$ for CL-CMRAC and FE-CMRAC are manipulated so that the both methods have the similar minimum eigenvalue as the proposed method, while the condition number $\kappa(F_k)$ and $h_k$ of the proposed method are much smaller than the other methods. As we can see in Fig. 1, this manipulation highlights that the low $\kappa(F_k)$ and $h_k$ can induce the low UUB, which is expected from Theorem 1. However, it should be noted that as a strategy of constructing the composite term affects the closed-loop system responses and the observed data $(\xi, y)$, it is difficult to conclude that low $\kappa(F_k)$ and $h_k$ always cause low UUB of CMRAC systems. Rather, the contribution of this study is that the proposed method can theoretically control the characteristics of the composite term as shown in Fig. 2, where the proposed method satisfies the condition in (7) that keeps $\kappa(F_k)$ and $h_k$ bounded while $\lambda_1(F_k)$ monotonically increases as predicted by Theorems 2 and 3.

## VII. Conclusion

In this technical note, we proposed a theoretically guaranteed method to manipulate the characteristics of CMRAC composite term that affects the UUB of system and the condition number of estimate. The set of weights for the composite term to have the desired properties was given by rank-one update technique. Also, as an example, a convex optimization problem was provided, which maximizes the minimum effective eigenvalue while keeping other conditions,

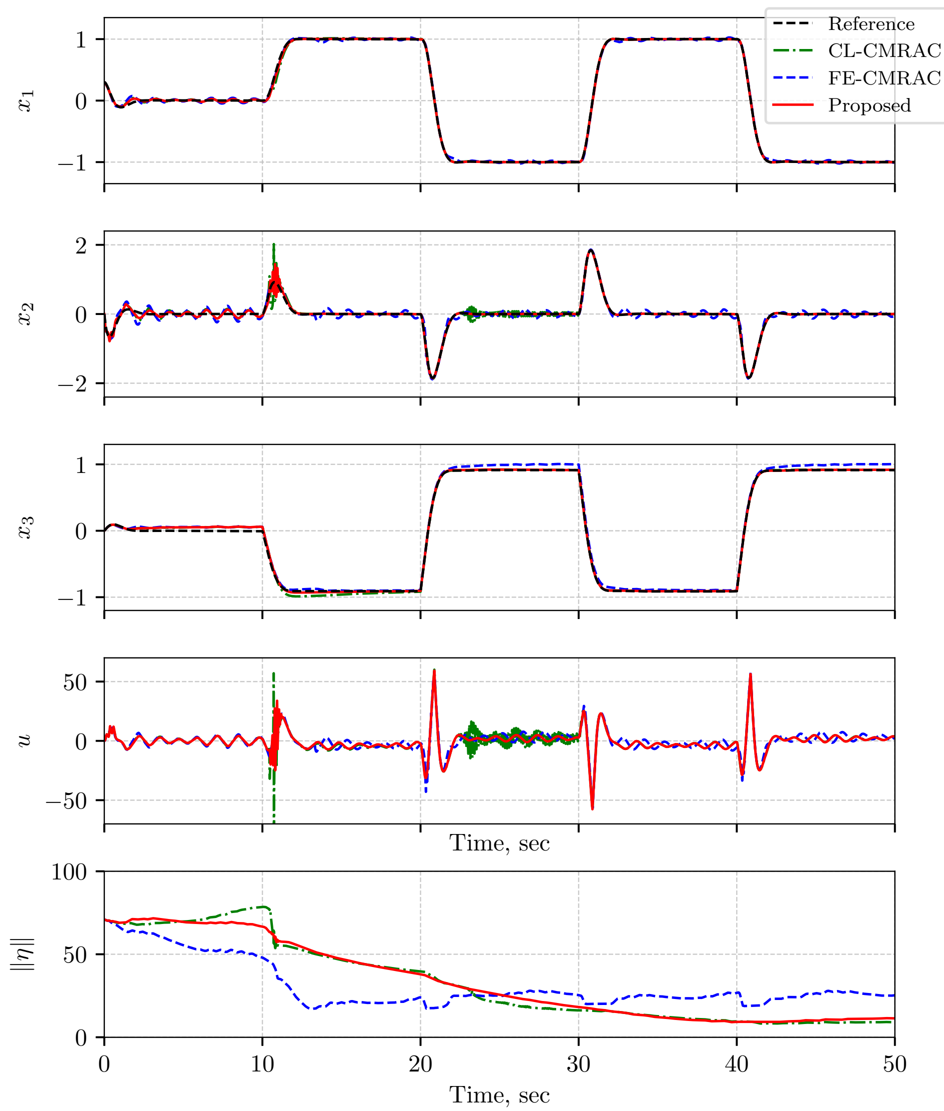


Fig. 1. The state, control input, and tracking error responses.

and an illustrative numerical simulation was also presented. Future research should consider the extensions of the proposed method to output feedback.

## Acknowledgment

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2019R1A2C208394612).

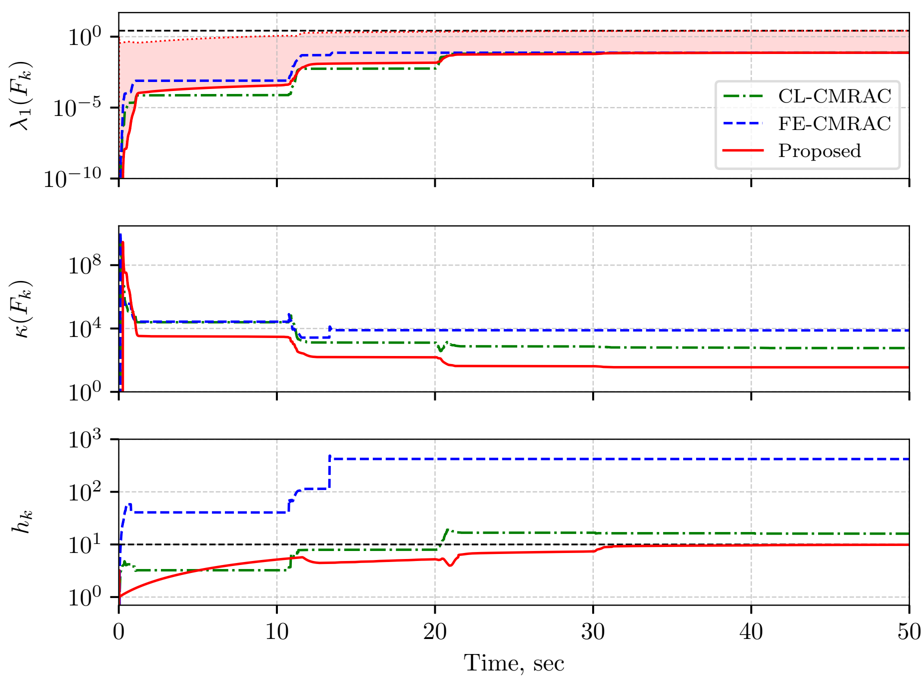


Fig. 2. The composite term characteristics of each method. The top graph shows the history of minimum eigenvalues where the red filled area indicates the range between minimum and maximum eigenvalues of the proposed method. The black dashed lines in the top and the bottom graphs represent the bounds of the maximum eigenvalue and $h_k$ set for the proposed method, respectively.

## References

[1] J.-J. E. Slotine and W. Li, "Composite adaptive control of robot manipulators," *Automatica*, vol. 25, no. 4, pp. 509–519, 1989.

[2] M. A. Duarte and K. S. Narendra, "Combined direct and indirect approach to adaptive control," *IEEE Trans. Autom. Control*, vol. 34, no. 10, pp. 1071–1075, 1989.

[3] M. K. Ciliz, "Combined direct and indirect adaptive control of robot manipulators using multiple models," *Advanced Robotics*, vol. 20, no. 4, pp. 483–497, 2006.

[4] E. Lavretsky, "Combined/composite model reference adaptive control," *IEEE Trans. Autom. Control*, vol. 54, no. 11, pp. 2692–2697, 2009.

[5] E. Lavretsky, R. Gadient, and I. M. Gregory, "Predictor-based model reference adaptive control," *J. Guid. Control Dyn.*, vol. 33, no. 4, pp. 1195–1201, 2010.

[6] G. Chowdhary and E. Johnson, "A singular value maximizing data recording algorithm for concurrent learning," in *Proceedings of the 2011 American Control Conference*, San Francisco, CA, Jun. 2011.

[7] G. Chowdhary, M. Mühlegg, and E. Johnson, "Exponential parameter and tracking error convergence guarantees for adaptive controllers without persistency of excitation," *Int. J. Control*, vol. 87, no. 8, pp. 1583–1603, 2014.

[8] Y. Pan and H. Yu, "Composite learning from adaptive dynamic surface control," *IEEE Trans. Autom. Control*, vol. 61, no. 9, pp. 2603–2609, 2016.

[9] Y. Pan and H. Yu, "Composite learning robot control with guaranteed parameter convergence," *Automatica*, vol. 89, pp. 398–406, 2018.

[10] N. Cho, H. Shin, Y. Kim, and A. Tsourdos, "Composite model reference adaptive control with parameter convergence under finite excitation," *IEEE Trans. Autom. Control*, vol. 63, no. 3, pp. 811–818, 2018.

[11] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, Upper Saddle River, NJ, 2002.

[12] B. Anderson, "Exponential stability of linear equations arising in adaptive identification," *IEEE Trans. Autom. Control*, vol. 22, no. 1, pp. 83–88, 1977.

[13] S. Boyd and S. Sastry, "Necessary and sufficient conditions for parameter convergence in adaptive control," *Automatica*, vol. 22, no. 6, pp. 629–639, 1986.

[14] E. Lavretsky and K. Wise, *Robust and Adaptive Control: With Aerospace Applications*, ser. Advanced Textbooks in Control and Signal Processing. Springer-Verlag, London, 2013.

[15] D. A. Belsley, E. Kuh, and R. E. Welsch, *Regression diagnostics: identifying influential data and sources of collinearity*, ser. Wiley series in probability and mathematical statistics. Wiley, NY, 1980.

[16] I. C. F. Ipsen and B. Nadler, "Refined perturbation bounds for eigenvalues of hermitian and non-hermitian matrices," *SIAM J. Matrix Anal. Appl.*, vol. 31, no. 1, pp. 40–53, 2009.

2022-08-09

# Data-efficient active weighting algorithm for composite adaptive control systems

Kim, Seong-hun

IEEE

Kim S-H, Lee H, Cho N, Kim Y. (2023) Data-efficient active weighting algorithm for composite adaptive control systems. IEEE Transactions on Automatic Control, Volume 68, Issue 5, May 2023, pp. 3086-3090

https://doi.org/10.1109/TAC.2022.3197702