Extremum seeking-based perfect adaptive tracking of non-PE references despite nonvanishing variance of perturbation

Miloje S. Radenković a,*, Miroslav Krštić b

a Department of Electrical Engineering, University of Colorado Denver, Denver, CO 801217, USA
b Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, CA 92093–0411, USA

1. Introduction

1.1. Adaptive control of systems with unknown control directions

It is well known that all well behaved adaptive controllers require information about the system control direction. The control direction is determined by the sign of the system’s high-frequency gain. When the sign of this gain is unknown, adaptive system can exhibit numerical instability due to the loss of stabilizability of the parameter estimates. Design of adaptive controllers in case of unknown control directions is a challenging problem. Inspired by the conjecture and questions raised by Morse (1983), Nussbaum (1983) proved that the knowledge of the sign of the high-frequency gain is not necessary requirement for adaptive stabilization. In this approach adaptive controller employs a mechanism for inverting the sign of the control signal if the system states start to grow without bound. Generalization of this result to higher order systems was presented in Mudgett and Morse (1985), and was further advanced in hundreds of subsequent publications (see for example a survey paper Ilchmann (1991)). Without relying on the Nussbaum approach, in this paper we develop a novel algorithm for reference tracking and adaptive stabilization of unstable systems with unknown control directions via perturbation-demodulation ES loop.

1.2. Extremum seeking for unstable plants

The perturbation-demodulation based ES scheme is convenient for finding and in real-time maintaining the optimizing value of an unknown input/output map. In this approach the estimate of the gradient of the nonlinear map is obtained by adding a perturbation signal to the optimizer estimates, and subsequently demodulating the observed output. The original intent of this method was to find optimal operating point of stable plants by automated tuning of system parameters. In the 1950s and 1960s this line of research went by names of extremal control, extremum regulation, hill-climbing regulation, etc. (Meerkov, 1967; Morosanov, 1957; Ostrovskii, 1957; Roberts, 1965). The absence of rigorous theory and difficulties in implementing ES controllers, in the 1970s lead to a decline in interest for this topic. The first stability analysis of an ES algorithm was published in 2000, and it is based on the method of averaging and perturbation theory (Krstic, 2000; Krstic & Wang, 2000; Wang & Krstic, 2000). Afterwards a large number of results covering various ES control topics have appeared (see for example the survey paper Tan, Moase, Manzie, Nešić, and Mareels (2010)). Since in this paper we are considering adaptive control of unstable systems, we will not review results related to ES optimization methods for stable plants. We only comment on contributions to ES based stabilization of unstable systems. A first use of ES loop to unstable plants is considered in Zhang, Siranosian,
1.4. Notation

The results reported in Moase and Manzie (2012) allow for the input dynamics to be unstable. In Hagi and Ariyur (2011) authors consider a model reference control via ES in case of unknown control directions and partial knowledge of system parameters. It is demonstrated that the tracking error converges globally to an $\epsilon(1/\omega)$ neighborhood of the origin, where $\omega$ is the frequency of the perturbation signal. A finite-time horizon ES based control for unknown and unstable linear discrete time systems is considered in Frihauf, Krstic, and Basar (2013). It is proved that a certain quadratic cost function converges locally, exponentially to some neighborhood of its optimal value. The first systematic design of ES control for unstable plants, along with a novel method to handle unknown control directions is presented in Scheinker and Krstic (2013). It is shown that the proposed controller provides semi-global exponential practical stability, while the system states converge to perturbation frequency dependent $\epsilon(1/\omega)$ proximity of the origin. Compared to Mudgett and Morse (1985), the above controller is robust to external disturbances and sign changes of the high-frequency gain. Our present work is related to Radenkovic and Krstic (2017), where the authors consider the problem of adaptive stabilization of possibly unstable linear discrete time systems with unknown control directions via ES. There it is assumed that: (i) the reference signal is persistently exciting (PE); (ii) the open loop system is irreducible; (iii) perturbation and demodulation signal are martingale difference sequences (m.d.s) whose variances tend to zero, and (iv) the employed parameter estimator has a vanishing gain sequence. It is proved that (a.s.) the tracking error converges to zero. Conditions (i)–(iii) are crucial for obtaining this result.

1.3. Contribution of this paper

This paper solves the problem of discrete time adaptive stabilization and optimal tracking in case of unknown control directions, by devising ES based algorithm with a non-vanishing perturbation. It is well known that the energy of perturbation signal affects the proximity within which the extremum point can be reached (Frihauf et al., 2013; Hagi & Ariyur, 2011; Krstic, 2000; Krstic & Wang, 2000; Moase & Manzie, 2012; Scheinker & Krstic, 2013). This proximity is on the order of the perturbation energy (see for example Theorem 2 and Corollary 2 in Frihauf et al., 2013). In this paper it is proved that (a.s.) the output tracking error converges to zero despite the non-vanishing (bounded from below) variance/energy of the perturbation signal. Similar problem has been considered in Radenkovic and Krstic (2017) under different assumptions and by using a different algorithm. The algorithm in Radenkovic and Krstic (2017) is a long memory algorithm whose gain converges to zero as time tends to infinity. The algorithm proposed in this paper has a non-vanishing gain, hence a short memory algorithm. In contrast to Radenkovic and Krstic (2017) in this paper: (1) we do not assume that the reference signal is persistently exciting; (2) open loop system is not required to be irreducible, and (3) the proposed controller enables to employ probing and demodulation sequences with non-vanishing variances. The results in Radenkovic and Krstic (2017) require that these variances converge to zeros, as a consequence of which the resulting scheme is a long memory algorithm. Because in this paper the variance of demodulation signal does not tend to zero, the resulting parameter estimator has a non-vanishing gain. In addition to obtaining asymptotically zero tracking error, we show that (a.s.) input and output signals remain bounded, while the parameter estimates are convergent sequences. This result is global in the sense that it holds for all initial conditions.

1.4. Notation

The subscript $T$ denotes the transpose of the matrix; $\mathbb{R}$ is the set of real numbers; $\mathbb{R}^n$ is the set of n-dimensional vectors with real entries; $\|x\|$ denotes the Euclidean norm of the vector $x$; $\text{sgn}(b)$ is the sign function of a $b \in \mathbb{R}$; $l_2$ denotes the normed vector space of sequences $\{e(k)\}$, $k \geq 0$ that are square summable, i.e., $\sum_{k=0}^{\infty} |e(k)|^2 < \infty$; the $l_{\infty}$ sequence space is defined as $\{e(k)\} \in \mathbb{R} : \sup_{k \geq 0} |e(k)| < \infty$.

2. Problem statement

Consider the following discrete-time system

$$A(q^{-1})y(k+1) = b_0 B(q^{-1})u(k), \quad b_0 \neq 0$$

(1)

where $y(k) \in \mathbb{R}$ and $u(k) \in \mathbb{R}$ are output and input sequences, respectively, $k = 0, 1, 2, \ldots$, is the discrete time index, and $q^{-1}$ is a unit delay operator. Polynomials $A(q^{-1})$ and $B(q^{-1})$ are given by

$$A(q^{-1}) = 1 + \sum_{i=1}^{l} a_i q^{-i}, \quad B(q^{-1}) = 1 + \sum_{i=1}^{l} b_i q^{-i}, \quad L > 0.$$ (2)

It is assumed that the parameters $b_0, a_i, b_i, i = 1, \ldots, L$, are unknown. The parameter $b_0$ is often referred to as the high frequency gain, or the ‘control direction’. In the above, $L$ represents an upper bound (known to a designer) on the unknown actual system order $L^* \leq L$. Therefore in (2) we have $a_i = 0, b_i = 0$ for $i = L^* + 1, \ldots, L$. The aim of this paper is to find the input sequence $u(k)$ so that for a given reference trajectory $y^\ast(k)$, the tracking error $e(k+1) = y(k+1) - y^\ast(k+1)$ satisfies $\{e(k)\} \in l_2$. It is assumed that $y^\ast(k), k \geq 0$, is generated by the following model

$$D(q^{-1})y^\ast(k+1) = 0, \quad y^\ast(0) \neq 0, \quad k \geq 0,$$ (3)

where $D(q^{-1})$ is known polynomial defined by

$$D(q^{-1}) = 1 + d_1 q^{-1} + \cdots + d_N q^{-N}, \quad N > 0.$$ (4)

For example if $D(q^{-1}) = 1 - q^{-1}$, then $y^\ast(k) = y^\ast(0), \forall k \geq 0$. If $D(q^{-1}) = 1 + d_1 q^{-1} + \cdots + d_1 q^{-N}$, then $y^\ast(k)$ is a cosine sequence with frequency $\omega_0$. Observe that in the above examples the zeros of $D(q^{-1})$ satisfy $|\omega_0| = 1$. To motivate the development of the control algorithm, we define the following signals

$$e_r(k) := D(q^{-1})e(k), \quad u_r(k) := D(q^{-1})u(k).$$ (5)

Then from (1) and (3) it follows that

$$A(q^{-1})e_r(k+1) = b_0 B(q^{-1})u_r(k),$$ (6)

or

$$e_r(k+1) = -q \left[A(q^{-1}) - 1\right] e_r(k) + b_0 B(q^{-1})u_r(k).$$ (7)

where $q$ is the forward shift operator, e.g., $q e_r(k) := e_r(k+1)$. Similarly from (3) and (6) we can write

$$e_r(k+1) = e_r(k+1) + D_1(q^{-1})e_r(k),$$ (8)

where $D_1(q^{-1}) = q \left[D(q^{-1}) - 1\right]$. After substituting (9) on the LHS of (8) one can obtain

$$e_r(k+1) = b_0 \theta^T \psi_r(k) + b_0 u_r(k),$$ (10)

where $\theta \in \mathbb{R}^{d+1}$ is given by

$$\theta^T = \begin{bmatrix} -1/b_0, & -a_1, & \ldots, & -a_i, & b_1, & \ldots, & b_L \end{bmatrix},$$ (11)

with $a_i = a_i/b_0, 1 \leq i \leq L$ and

$$\psi_r(k) = \left[D_1(q^{-1}) e_r(k), e_r(k), \ldots, e_r(k-L+1), u_r(k-L+1), \ldots, u_r(k-L)\right]$$ (12)

where $e_r(k)$ and $u_r(k)$ are defined by (6). From (10) it is obvious that the control law $u_r(k) = -\theta^T \psi_r(k)$ gives $e(k+1) = 0, \forall k \geq L,$
provided that \(1/B(\xi^{-1})\) is a stable operator. In case the parameter vector \(\theta\) is unknown we adopt the certainty-equivalence principle, and the control input is calculated by replacing \(\theta\) with \(\theta_p(k)\) obtained based on the observations of \(e(i)\) and \(\varphi(i-1), 0 \leq i \leq k\). For generating \(\theta_p(k)\) we propose an ES based algorithm whose functional block diagram is shown in Fig. 1, where \(\hat{\theta}(k)\) is the estimate of the unknown parameter vector \(\theta\), \(q\) is the forward shift operator \((q \hat{\theta}(k) = \hat{\theta}(k+1))\), and \(1/(q - 1)\) represents discrete time integrator. The scheme depicted there adjust \(\hat{\theta}(k)\) so that the cost function \(J(\hat{\theta}) = e(k+1)^2\) achieves its minimum. The scheme actually represents perturbation-demodulation loop, where \(p(k)\) (see Fig. 1) is the perturbation signal added to the estimates \(\hat{\theta}(k)\), and \(d(k)\) is the demodulation signal. While passing through the nonlinear function \(J(\hat{\theta})\) the amplitude of \(p(k)\) is modulated by the gradient of \(J(\hat{\theta})\). The observed quantity \(e(k+1)^2\) is then demodulated, thereby producing the estimate of the gradient direction. We now define signals \(p(k)\) and \(d(k)\). Let

\[
r(k) = r(k-1) + \|\varphi(k)\|^2, \quad r(0) > 1,
\]

where \(\varphi(k)\) is given by (12). Define

\[
h(k) := H(k)p(k)[\log r(k)]^{1-\varepsilon_1}/2, \quad 0 < \varepsilon_1 < 1,
\]

where

\[
H(k) := 1 + \max_{0 \leq \varepsilon \leq k} \|\hat{\theta}(\varepsilon)\|^2.
\]

We choose the following perturbation sequence

\[
p(k) = g(k)w(k+1), \quad k \geq 1,
\]

where \(\{w(k)\} \in \mathbb{R}^{2L+1}\) is a random sequence specified below (see Assumption 3), and

\[
g(k) = 1/|\log r(k)|^{1-\varepsilon_1}/2,
\]

with the same \(r(k)\) and \(\varepsilon_1\) as in (13) and (14), respectively. The demodulation sequence is selected as

\[
d(k) = \mu \frac{w(k+1)}{h(k)}, \quad \mu > 0, \quad k \geq 1,
\]

where \(h(k)\) is defined by (14). The input signal is given by \(u_i(k) = -\theta_p(k)\varphi(k)\) where \(\theta_p(k)\) is the perturbed value of \(\theta(k)\) (see Fig. 1). Thus

\[
u_i(k) = -\left(\hat{\theta}(k) + p(k)\right)^T\varphi(k),
\]

where \(\hat{\theta}(k)\) is the estimate of the unknown \(\theta\), and \(p(k)\) is given by (16). From Fig. 1 and (18) we can write \(\forall k \geq 0\)

\[
\hat{\theta}(k+1) = \hat{\theta}(k) - \mu \frac{w(k+1)\varphi(k+1)^2}{h(k)}, \quad \mu > 0
\]

where \(e(k+1)\) and \(h(k)\) are defined by (3) and (14), respectively. The initial \(\hat{\theta}(0)\) is arbitrary. Note that the actual \(u_i(k), k \geq 1\), is calculated from (19) and (6) by using \(u_i(k) = u_i(k) - q[D(q^{-1}) - 1]u_i(k-1)\). The equation describing the closed-loop system dynamics plays an important role in our analysis and it is obtained by substituting (16) in (19), and then (19) in (10), i.e.,

\[
e(k+1) = -b_d(x(k) - b_0g(k)w(k+1)^T\varphi(k),
\]

where

\[
x(k) := \hat{\theta}(k)^T\varphi(k), \quad \hat{\theta}(k) := \hat{\theta}(k) - \theta.
\]

Next we introduce assumptions about the system (1), the reference signal \(4\), and the perturbation/demodulation sequence \(\{u(k)\}\).

**Assumption 1** (Minimum Phase System). Polynomial \(B(\xi^{-1})\) has zeros strictly inside the unit circle.

**Assumption 2** (Concerning the Reference Model (4)). The zeros of \(D(\xi^{-1})\) are on the unit circle and there are no repeated zeros.

**Assumption 3** (Concerning \(\{w(k)\}\)). The sequence \(\{w(k)\}\) is defined on the underlying probability space \((\Omega, \mathcal{F}, \mathcal{P})\). Let \(w(k) = w_1(k), \ldots, w_{2L+1}(k)\). Then \(\{w_i(k)\}, 1 \leq i \leq 2L+1\), is a martingale difference sequence (m.d.s) with respect to an increasing sequence of \(\sigma\)-fields \(\mathcal{F}_k, (\mathcal{F}_k \subseteq \mathcal{F})\) generated by \(\{w_0, w_1, \ldots, w_k\}\), i.e., \(w_i(k)\) is \(\mathcal{F}_k\)-measurable, and \(E[w_i(k+1)|\mathcal{F}_k] = 0, (a.s.), \forall k \geq 0\). It is assumed that the elements of \(w(k)\) are mutually independent processes. In addition the following holds (a.s.): \(E[w_i(k+1)^2|\mathcal{F}_k] = \sigma^2, E[w_i(k+1)^2|\mathcal{F}_k] = 0, i = 1, \ldots, 2L + 1, \) and \(\|w(k)\| \leq \epsilon_k < \infty, \forall k \geq 0\).

**Comments about the assumptions:**

(i) Assumption 1 is often used in adaptive control theory, and it states that (1) is a minimum phase system, implying that the inverse plant characterized by \(A(q^{-1})/B(q^{-1})\) is stable.

(ii) Since in this paper we are dealing with bounded reference signals \(y'(k)\), it must be assumed that \(D(q^{-1})\) (see (4)) has no zeros outside of the unit circle, and the zeros on the unit circle must be simple. The zeros of \(D(q^{-1})\) located strictly inside the unit circle create exponentially decaying modes, thus not affecting the asymptotic behavior of the reference \(y'(k), k \geq 0\). Hence it is reasonable to assume that Assumption 2 holds.

(iii) The sequence \(\{w(k)\}\) characterized by Assumption 3 represents a probabilistic model of a uniformly distributed “white” noise process. Since \(\{w(k)\}\) is selected by the designer Assumption 3 is not a restrictive constraint.

**Remark 4** (Selection of the Sequences \(\{g(k)\}\) and \(\{h(k)\}\) in (14), (17) and (20)). At the core of the global stability analysis presented in Section 4, is the construction of a positive near supermartingale \(\{V(k), \mathcal{F}_k\}\) to which we then apply the Martingale Convergence Theorem (MCT) (see p. 848 in Goodwin, Ramadge, and Caines [1981]), to conclude that (26) and (47) hold. This analysis reveals that \(g(k)\) and \(h(k)\) (or \(H(k)\) and \(r(k)\)) are selected (designed) so that the following holds:

(i) \(\{h(k)\}\) and \(\{1/g(k)\}\) must be monotonically nondecreasing so that \(h(k) = h(k)\) satisfies \(h(k) \geq h(k-1), \forall k \geq 1\). Monotonicity of \(h(k)\) is needed for deriving (46) from (36) and (45);

(ii) \(\sum_{k=0}^{\infty} \left(1 + \|\hat{\theta}(k)\|^2\right)^2 \|\varphi(k)\|^4/h(k)^2 < \infty\), which in turn requires that \(h(k)\) must contain factors \(H(k)\) (see (15) and its role in deriving (A.8) from (A.7) in the Appendix), and \(r(k)\) satisfying (A.9) in the Appendix and \(\|\psi(k)\|^2/r(k) < \infty\) (see its use in deriving (A.8) from (A.7)). The aforementioned condition is part of the proof of (24) stating that the 3rd term on the RHS of (46) is infinitely summable. This fact facilitates the application of MCT to (46) to obtain (26) and (47), and subsequently (27).
where \( e \) and constant \( C \) is dropped from \{ \( F \) \}.

It should be pointed out that \( g(k) \) and \( \mu/h(k) \) define the variances of the perturbation, and demodulation sequences, respectively (see (16) and (18)). The variable \( \mu/h(k) \) represents the gain of the parameter estimator (20), and the evolution of \( h(k) \) determines whether (20) has vanishing or non-vanishing gain. At first glance (13) and (14) may give impression that \( h(k) \to \infty \) as \( k \to \infty \), thus resulting in the vanishing algorithm gain. To continue our discussion along those lines let us assume for a moment that \( \theta(k) = \theta, \forall k \geq 0 \). Then (21) implies \( e(k+1) = -b_0 g(k) w(k+1) \psi(k) \), from where on the surface may seem that the tracking error \( e(k+1) \) will tend to zero only if \( g(k) \to 0 \) (or \( r(k) \to \infty \)) as \( k \to \infty \) (see (17)). Theorem 7 proves that this is not so, i.e., it is shown that \( \theta(k) \in l_o, \forall k \in l_o, \) and \( r(k) \in l_o \). Consequently \( h(k) \) will be a bounded sequence (see (13) and (14)) implying that the estimator (20) has non-vanishing gain and \( p(k) \) and \( d(k) \) have nonzero variances.

### 3. Technical results

Throughout this paper we use nonnegative numbers \( c_i, i = 1, 2, \ldots \), to denote certain upper bounds whose specific values are unimportant. Numbers \( c_i, i \geq 1 \) are independent of the time indexes and may depend on a sample point \( \omega \in \Omega \), where \( \Omega \) is the sample space. For the sake of notational simplicity the argument \( \omega \) is dropped from \( c_i(\omega) \) and we simply write \( c_i, i \geq 1 \). The following technical results are used in the stability and convergence analysis of the proposed algorithm.

**Lemma 5.** Let Assumptions 1 and 3 hold. Then for some positive constant \( C_e \) and \( c_1 \), we have

\[
\frac{r(k)}{C_e} \leq \sum_{i=1}^{k} e(i)^2 + c_1, \forall k \geq 0, \quad (23)
\]

and

\[
l_1 := \sum_{k=1}^{\infty} \left[ \frac{e(k+1)^4}{h(k)^2} \right] < \infty, \quad (a.s.), \quad (24)
\]

where \( e(k) \) and \( h(k) \) are defined by (3) and (14), respectively.

**Proof.** The proof is given in the Appendix.

**Lemma 6.** Let Assumption 3 hold. Then

\[
E \left\{ \frac{\theta(k)^T w(k+1)}{h(k)} g(k)^2 \left[ w(k+1)^T \psi(k) \right] \right\}_{\mathcal{F}_k} = 0, \quad (a.s.) \quad (25)
\]

where \( \psi(k), h(k), g(k) \) and \( \theta(k) \) are defined by (12), (14), (17), and (22), respectively.

**Proof.** The following arguments justify statement (25). The LHS of (25) contains terms of the form \( E[\sigma_{j}(\omega) w(k+1) w(k+1)^T \psi(k)] \), \( 1 \leq j, i, n \leq 2L+1 \), where \( \sigma_{j}(\omega) \) is independent of \( g(k), h(k) \) and the elements of \( \psi(k) \) and \( \theta(k) \). Hence \( \sigma_{j}(\omega) \) is a \( \mathcal{F}_n \)-measurable variable. Therefore (25) follows from the fact that by Assumption 3, \( E \left\{ w(k+1)^T \psi(k) \right\}_{\mathcal{F}_k} = 0, \quad (a.s.) \) and the elements of \{w(k)\} are mutually independent m.d.s.

### 4. Global stability, tracking performance and convergence of the estimates \( \{\theta(k)\} \)

In this section we examine the nonlinear dynamics defined by (20) and (21), and the evolution of parameter estimates \( \hat{\theta}(k, \omega), k \geq 0 \) for a fixed sample point \( \omega \in \Omega \). To shorten the notation the argument \( \omega \) is omitted, and instead of \( \hat{\theta}(k, \omega) \) we write \( \hat{\theta}(k) \). We prove that the proposed algorithm guarantees \( e(k) \in l_2, \psi(k) \in l_o, \) and \( \hat{\theta}(k) \) is a convergent sequence. The above claims are global in the sense that they hold for all initial conditions \( \hat{\theta}(0) \) and \( \psi(0) \).

**Theorem 7.** Let Assumptions 1–3 hold. Then the algorithm given by (19) and (20) provides

\[
\| \hat{\theta}(k) \| \leq v_1 < \infty, \forall k \geq 0, \quad (a.s.), \quad (26)
\]

where \( \hat{\theta}(k) \) is defined by (22), and \( v_1 \) depends on \( \omega \in \Omega \).

\[
E \left\{ e(k)^2 \right\} < \infty, \quad (a.s.), \quad (27)
\]

\[
\| \psi(k) \| \leq \| \psi(0) \|, \quad (a.s.), \quad (28)
\]

\[
\bar{\theta} := \lim_{k \to \infty} \hat{\theta}(k) \text{ exists,} \quad (a.s.), \quad (29)
\]

for all finite initial conditions \( \hat{\theta}(0) \) and \( \psi(0) \).

**Proof.** We first establish that for a fixed \( \omega \in \Omega \) the parameter estimate \( \hat{\theta}(k, \omega) \) is (a.s.) bounded sequence. We do this by constructing a suitable stochastic Lyapunov function corresponding to (20), and invoking the martingale convergence theorem. Then (26) is used to prove (27). Statements (28) and (29) are derived from (26), (20) and (27).

**Proof of (26).** After subtracting \( \theta \) from both sides of (20) and squaring the obtained equation one can derive

\[
\| \hat{\theta}(k+1) \|^2 = \| \hat{\theta}(k) \|^2 - 2\mu \frac{\hat{\theta}(k)^T w(k+1) e(k+1)}{h(k)} + \mu^2 \| w(k+1) \|^2 e(k+1)^2 \quad (h(k)^2) \quad (30)
\]

where \( \hat{\theta}(k) \) is defined by (22). Observe that from (21) we have

\[
E \left[ e(k+1)^2 \right] = b_0^2 x(k)^2 + 2 b_0 x(k) g(k) w(k+1)^2 \psi(k) + b_0^2 g(k)^2 \left( w(k+1)^2 \psi(k) \right)^2. \quad (31)
\]

By substituting (31) in the 2nd term on the RHS of (30), and taking into account that in the 3rd term on the RHS of (30) \( \| w(k+1) \| \leq c_w, \quad (a.s.) \) (see Assumption 3) one can obtain that \( \forall k \geq 0, \)

\[
\| \hat{\theta}(k+1) \|^2 \leq \| \hat{\theta}(k) \|^2 - 2\mu \frac{\hat{\theta}(k)^T w(k+1) b_0}{h(k)} x(k)^2 + \psi(k) + g(k)^2 \left( w(k+1)^2 \psi(k) \right)^2 + \mu^2 c_w e(k+1)^4 \quad (h(k)^2) \quad (a.s.) \quad (32)
\]

We next calculate the conditional expectation \( E(\cdot | \mathcal{F}_k) \) of the 2nd term on the RHS of (32). Owing to the fact that \( \psi(k), h(k), g(k), \) and \( \theta(k) \) are \( \mathcal{F}_n \)-measurable variables, Assumption 3 implies

\[
E \left\{ \frac{\hat{\theta}(k)^T w(k+1) x(k)^2}{h(k)} \right\}_{\mathcal{F}_k} = 0, \quad (a.s.) \quad (33)
\]

\[
E \left\{ \frac{\hat{\theta}(k)^T w(k+1) x(k)^2}{h(k)} \right\}_{\mathcal{F}_k} = 0, \quad (a.s.)
\]
Taking into account that by (33) 

\[
\sum_{i=0}^{k} \|x(i)\|^2 \leq r(k) \leq (C_r + \varepsilon_3) \sum_{i=0}^{k} e(i)^2, \quad (a.s.),
\]

(42)

for some \( \varepsilon_3 > 0, \varepsilon_3 \ll 1 \). Note that in (23) the constant \( c_1 \) is dominated by \( \sum_{i=0}^{k} e(i)^2 \). By using (42) in (41) we obtain

\[
\sum_{i=0}^{k} \|x(i)\|^2 \geq \rho_1 \sum_{i=0}^{k} e(i)^2, \quad k \geq k_0, \quad (a.s.),
\]

(43)

where \( \rho_1 = (1/2b_0^2) - \varepsilon_2 (C_r + \varepsilon_3) > 0 \). Here it is assumed that in (41) \( \varepsilon_2 \) is selected so that \( \rho_1 > 0 \). Obviously from (43) it follows that there exists a positive constant \( c_{13} \) such that

\[
W(k) := 4\mu b_0^2 \sigma_0^2 \sum_{i=0}^{k} [x(i)^2 - \rho_1 e(i + 1)^2] + c_{13} \geq 0, \quad (a.s.),
\]

(44)

for all \( k \geq 0 \). Define

\[
V(k + 1) := \|\hat{\theta}(k + 1)\|^2 + \frac{W(k)}{\hat{h}(k)}.
\]

Then by using the monotonicity of \( \hat{h}(k) \) (see (13) and (35)) from (36) and (44) one can derive

\[
E \left\{ V(k + 1) \big| \mathcal{F}_k \right\} \leq V(k) - \rho_2 \frac{e(k + 1)^2}{\hat{h}(k)} + \mu^2 c_0^2 E \left\{ \frac{e(k + 1)^4}{\hat{h}(k)^2} \big| \mathcal{F}_k \right\}, \quad (a.s.),
\]

(46)

where \( \rho_2 = 4\mu b_0^2 \sigma_0^2 \rho_1 \). Similarly as in deriving (26) (see the analysis after (32)), by taking into account (24), (46) and the Martingale Convergence Theorem (Goodwin et al., 1981, p. 848), it follows that \( V(k) \) and \( \sum_{i=0}^{k} e(i + 1)^2/\hat{h}(i) \) converge to finite random variables, e.g., \( V(k) \rightarrow v_2 < \infty, (a.s.) \), as \( k \rightarrow \infty \), and

\[
\lim_{k \rightarrow \infty} \sum_{i=0}^{k} \frac{e(i + 1)^2}{\hat{h}(i)} = v_3 < \infty, \quad (a.s.),
\]

(47)

where \( v_2 \) and \( v_3 \) depend on \( \omega \in \Omega \). Thus (39) is proved.

**Proof of part 2 of (27).** We now examine \( \hat{h}(i) \) in (47) (see definition (35)). Define \( r_\varepsilon(k) \) as follows

\[
r_\varepsilon(k) := r_\varepsilon(k - 1) + e(k)^2, \quad r_\varepsilon(0) = c_1/C_r,
\]

(48)

\( \forall k \geq 1 \), where \( c_1 \) and \( C_r \) are the same as in (23). Hence (23) yields

\[
r(k) \leq C_r r_\varepsilon(k), \quad \forall k \geq 0.
\]

(49)

On the other hand boundedness of \( \tilde{h}(k) \) (see (26) and (15)) imply \( H(k) \leq c_{14} < \infty, (a.s.) \), \( \forall k \geq 0 \), where \( c_{14} \) depends on \( \omega \in \Omega \). Thus, by using (48) and (49) in (35) gives \( \tilde{h}(k) \leq c_{14} C_r r_\varepsilon(k) \log (C_r t_r(k)), (a.s.) \), \( \forall k \geq 1 \). Because \( r(k) \rightarrow \infty, (a.s.) \) (see (38)), relation (49) leads to \( r_\varepsilon(k) \rightarrow \infty \) as \( k \rightarrow \infty \), (a.s.), implying that \( \log (C_r t_r(k)) \leq 2 \log r_\varepsilon(k), (a.s.) \). Based on the above discussion we can conclude that for some positive constant \( c_{15} \) the following holds, \( \tilde{h}(k) \leq c_{15} C_r r_\varepsilon(k), (a.s.) \), \( \forall k \geq 1 \), from where by (39) or (48) it follows that \( \forall k \geq 0 \),

\[
\sum_{i=0}^{k} \frac{e(i + 1)^2}{r_\varepsilon(i) \log r_\varepsilon(i)} \leq v_4 < \infty, \quad (a.s.),
\]

(50)
where $v_4$ depends on $\omega \in \Omega$. By virtue of the fact that $e(i+1)^2 = r_q(i+1) - r_q(i)$ (see (48)), from (50) one can derive $\forall k \geq 0$,

$$v_4 \geq \sum_{i=0}^{k} \frac{1}{r_q(i) \log r_q(i)} \int_{r_q(i)}^{r_q(i+1)} \frac{dz}{z \log z}$$

$$\geq \sum_{i=0}^{k} \int_{r_q(i)}^{r_q(i+1)} \frac{dz}{z \log z}$$

$$= \log \log r_q(k+1) - \log \log r_q(0), \quad (a.s.)$$  \tag{51}

or $r_q(k+1) \leq \exp\{\exp(v_4) \log r_q(0)\}, \forall k \geq 0, \quad (a.s.).$ Recall that the above conclusion is obtained under the assumption that (38) holds. Hence by (48) and (51) it follows that if $\lim_{k \to \infty} r_q(k) = \infty$ (see (38)), then $\{e(k)\} \in l_2$. If $\{r(k)\} \in l_\infty$ by (37) we have $\{e(k)\} \in l_2$. Thus (27) is proved.

Let us prove (28). By (4) and Assumption 2 it follows that $\{y^r(k)\} \in l_\infty$ from where (5) and (27) one concludes that $\{y(k)\} \in l_\infty$. Then the input sequence $\{u(k)\} \in l_\infty$ due to the fact that model (1) is a minimum phase system. Thus (28) is proved.

We next show (29). Backward iteration of (20) with respect to $\hat{\theta}(k)$ gives,

$$\hat{\theta}(k+1) = \hat{\theta}(0) - \mu \sum_{i=0}^{k} \frac{w(i+1)e(i+1)^2}{h(i)}, \quad \forall k \geq 0. \quad \tag{52}$$

Since by Assumption 3, $\|w(k)\| \leq c_w < \infty \quad (a.s.)$, and $h(k) > 0, \forall k \geq 0$ (see (14)), from (27) one obtains

$$\left| \sum_{i=0}^{k} \frac{w(i+1)e(i+1)^2}{h(i)} \right| \leq c_{16} < \infty, \quad \forall k \geq 0, \quad (a.s.) \quad \tag{53}$$

Hence, the 2nd term on the RHS of (52) is a partial sum of an absolutely convergent series, implying that (29) holds. It should be noted that in (29) $\theta = \hat{\theta}(\omega)$ depends on a sample point $\omega \in \Omega$. Thus the theorem is proved.

Remark 8. Observe that by (27), (23) and (13) we have $\|\psi(k)\| \in l_2, \quad (a.s.)$. Consequently $\psi(k)$ is not persistently exciting regardless of the frequency content of $\{y^r(k)\}$. Hence in the absence of persistent excitation (PE) the parameter estimation error $\hat{\theta}(k)$ (see (22)) will not converge to zero. It seems that the existence of a limit point $\check{\theta}$ (see (29)) is all we can say about the convergence of $\{\hat{\theta}(k)\}$.

Remark 9 (Comparison with the Results in Radenkovic & Krstic, 2017). In this paper it is presumed that the reference signal $y^r(k), k \geq 0$ is generated by a known model (4). The work in Radenkovic and Krstic (2017) does not assume existence of a reference model, however it supposes that the reference is persistently exciting. The above differences in assumptions related to $\{y^r(k)\}$, coupled with the differences between employed adaptive algorithms, give rise to different results. In Radenkovic and Krstic (2017) it is proved that (a.s.) the tracking error tends to zero, and the parameter estimates converge towards their true values, provided that respective perturbation signal has a vanishing variance. Since this paper does not impose PE condition on $\{y^r(k)\}$, we cannot establish consistency of $\hat{\theta}(k)$, i.e., $\theta(k) - \hat{\theta}(k) \to 0$ (a.s.) as $k \to \infty$.

On the other hand in contrast to Radenkovic and Krstic (2017), the variance of $p(k)$ in (16) does not converge to zero. This follows from the fact that by (27) and (23), $\{r(k)\} \in l_\infty, \quad (a.s.)$. Then from (16) and (17), and Assumption 3 we conclude that (a.s.), $E[p(k)^2] \geq p_0 > 0, \forall k \geq 0$. Furthermore by virtue of the fact that $\{r(k)\} \in l_\infty$ and $\{\theta(k)\} \in l_\infty$, from (14) and (15) it follows that the parameter estimator (20) has a non-vanishing gain. In Radenkovic and Krstic (2017) the estimator gain converges to zero.

5. Simulation example

In this example the ES based algorithm defined by (13), (14), (16), (19), and (20) is applied to the following system $\{1 - 0.7q^{-1} + 3.5q^{-2}\} y^r(k+1) = -2.5(1 + 0.5q^{-1}) u(k), k \geq 0$. The open loop system is unstable. Parameters of the system and the control direction (sign $(-2.5)$) are unknown. Perturbation and demodulation effects (see (16) and (18)) are determined by the vector sequence $\{w(k)\}$ whose elements are $\{w_i(k)\}, i = 1, \ldots, 2L+1$. Each $\{w_i(k)\}$ consists of independent and identically distributed (i.i.d) samples drawn from a uniform distribution with zero mean value and variance $\sigma^2_w = 0.45$. Elements of $\{w(k)\}$ are mutually independent processes. It should be pointed out that the above $\{w(k)\}$ is a vector martingale difference sequence. In (14) and (17) we set $\varepsilon_1 = 0.1$. The reference trajectory is generated by $\{1 - q^{-1}(1 + p_1 q^{-1} + q^{-2}) y^r(k+1) = 0, \quad p_1 = -2\cos(\pi/50), \text{ and } y^r(0) = 5.247, y^r(-1) = 5.645, y^r(-2) = 6.0469\}. \quad \tag{55}$

Obviously $y^r(k)$ is a superposition of a constant signal and a sine wave of discrete frequency equal to $\pi/50$ rad/sample. Fig. 2 displays the tracking error $e(k) = y(k) - y^r(k)$. Fig. 3 depicts the output trajectory $y(k)$. The evolution of the parameter estimates $\hat{\theta}(k), k \geq 0$ is shown in Fig. 4. The variance of the perturbation sequence $\{p(k)\}$ is determined by $g(k)$ (see (16)). The inverse of $g(k)$ is shown in Fig. 5. Evidently, the variance of $p(k), k \geq 0$ does not vanish to zero. Simulation experiments indicate that larger $\sigma_w$, and $\varepsilon_1$ closer to one, result in shorter transient response, while generating larger transient error $e(k)$.\vspace{4pt}
where $c_2$ and $c_3$ are positive constants. The role of $c_3$ is to capture the effect of nonzero initial conditions. Similarly from (6) we can write
\[
\sum_{i=0}^{k} e(i)^2 \leq c_4 \sum_{i=0}^{k} e(i)^2 + c_5,
\]
(A.2)
where $0 < c_4 < \infty$ and $0 \leq c_5 < \infty$. Then by using (A.1) and (A.2), from (12) one can obtain
\[
\sum_{i=0}^{k} \|\psi(k)\|^2 \leq c_6 \sum_{i=0}^{k} e(i)^2 + c_7,
\]
(A.3)
for some positive constant $c_6$, and $0 \leq c_7 < \infty$. Definition of $r(k)$ (see (13)) and (A.3) imply (23).

**Proof of (24).** Note that from (21) one can derive
\[
e(k+1)^2 \leq 2b_0^2 \left[ x(k)^2 + g(k)^2 \|w(k+1)\|^2 \|\psi(k)\|^2 \right].
\]
(A.4)
On the other hand (22) implies
\[
x(k)^2 \leq \|\hat{\theta}(k)\|^2 \|\psi(k)\|^2 \leq 2 \left( \|\hat{\theta}(k)\|^2 + \|\theta\|^2 \right) \|\psi(k)\|^2.
\]
(A.5)

Because $\|w(k)\| \leq c_w < \infty$, (a.s.) (see Assumption 3), and by (17) and (13) $g(k) \leq c_b < \infty$, $\forall k \geq 0$, from (A.4) and (A.5) one can conclude that
\[
e(k+1)^2 \leq c_9 \left( 1 + \|\hat{\theta}(k)\|^2 \right) \|\psi(k)\|^2, \text{ (a.s.)}
\]
(A.6)
where the positive constant $c_9$ may depend on a sample point $\omega \in \Omega$. Since by (A.6)
\[
e(k+1)^4 \leq c_{10} \left( 1 + \|\hat{\theta}(k)\|^2 \right)^2 \|\psi(k)\|^4, \text{ (a.s.)}
\]
(A.7)
for $k \geq 0$, from (14) we have that (a.s.) $\forall k \geq 0$
\[
E \left\{ \frac{e(k+1)^4}{h(k)^2} \bigg| \mathcal{F}_k \right\} \leq \frac{c_{10} \left( 1 + \|\hat{\theta}(k)\|^2 \right)^2 \|\psi(k)\|^2}{H(k)^2 r(k)} \times \frac{\|\psi(k)\|^2}{r(k) [\log r(k)]^{1+r_1}} \leq c_{10} \frac{\|\psi(k)\|^2}{r(k) [\log r(k)]^{1+r_1}}.
\]
(A.8)
Thus, by taking into account the definition of $I_1$ (see (24)) and exploiting monotonicity of $r(k)$ (see (13)), from (A.8) one can obtain
\[
I_1 \leq c_{10} \sum_{k=1}^{\infty} \frac{\|\psi(k)\|^2}{r(k) [\log r(k)]^{1+r_1}} \leq c_{10} \sum_{k=1}^{\infty} \frac{1}{r(k) [\log r(k)]^{1+r_1}} \int_{r(k-1)}^{r(k)} dz
\]
\[
\leq c_{10} \sum_{k=1}^{\infty} \int_{r(k-1)}^{r(k)} \frac{dz}{z^{\log z^{1+r_1}}} \leq c_{11} < \infty, \text{ (a.s.).}
\]
(A.9)
Thus the lemma is proved.

**References**


Kumar, P. R., & Varaiya, P. (1986). *Stochastic system: Estimation, identification, and adaptive control*. Prentice-Hall.


