ADAPTIVE CONTROL VIA A SIMPLE SWITCHING ALGORITHM*

JI FENG ZHANG† AND PETER E. CAINES‡

Abstract. In this paper we present an adaptive stabilization control for systems with unknown constant parameters and stochastic disturbances, which may be neither open-loop stable nor minimum phase. The ideas come from previous works [J. F. Zhang and H. F. Chen, Adaptive stabilization under the weakest condition, Proc. 31st Control and Design Conference, December 14–18, 1992, pp. 3620–3621, and H. F. Chen, Continuous-Time Stochastic Adaptive Control Stabilizing the System and Minimizing the Quadratic Loss Function, Tech. Report, Institute of Systems Science, Academia Sinica, Beijing, 1992], but here we not only simplify the construction procedure of an adaptive control but also reduce the computational load significantly, so that the adaptive control in this paper is more practical. Furthermore, parameter estimation is carried out in only a finite time period and, unlike previous work, the parameter estimates are generated by ordinary differential equations rather than stochastic differential equations.

Key words. adaptive control, parameter estimation, switching algorithm, continuous time, stochastic system

AMS subject classifications. 93C40, 93E15, 93E35

1. Introduction. The switching control strategies of Zhang and Chen [1991] and Chen [1992] show that an alternation of excitation and control regimes can yield stabilizing controls. The idea is that, if a certain prediction error test fails at a specified instant, then a signal which is (in the limit) persistently exciting is applied. On the other hand, if the test is passed, then a particular certainty equivalence control law using the current estimate is applied. This strategy has common-sense appeal, despite the fact that the laws are somewhat complex in their present form. It is shown in the analysis of these laws that eventually the prediction error tests must always be passed, and hence it is shown that the system “locks on” to an acceptable control law. In summary, the adaptive control algorithms used in Zhang and Chen [1992] and Chen [1992] are as follows:

Step A) Introduce an appropriate criterion to judge whether or not the parameter estimate is satisfactory (for instance, a prediction error criterion).
Step B) Apply an excitation signal to the system, and estimate the unknown parameters via a least-squares (or related) algorithm until a “satisfactory” estimate is obtained according to the criterion; and after this,
Step C) construct a control law via the previously obtained “satisfactory” parameter estimates and use this law to control the system until some “unsatisfactory” property appears according to the criterion; and then
Step D) repeat this procedure through Steps B) and C).
If no “unsatisfactory” property appears at some stage in Step C), then the designed adaptive control law is used forever.

It is worth noticing that in some previous works (i) one or both of the derivatives $dx_t$ and $dy_t$ of the system state $x$ and observation process $y$ are required to be measurable in the parameter estimation procedure (see, e.g., Caines [1992]; Chen [1992];

* Received by the editors November 25, 1992; accepted for publication (in revised form) October 7, 1994. This research was supported by Canadian NSERC grant A 1329.
† Canadian Institute for Advanced Research, and Institute of Systems Science, Academia Sinica, Beijing 100080, People’s Republic of China.
‡ Department of Electrical Engineering, McGill University, 3480 University Street, Montreal H3A 2A7, Canada.
Chen and Guo [1990]; Chen and Moore [1987]; Duncan and Pasik-Duncan [1990], [1991]; Gevers, Goodwin, and Wertz [1991]; Goodwin, et al. [1991]; Moore [1988]; Christopeit [1986]); (ii) the criteria used in steps A through C have to be verified at all time instants, which is an uncountable procedure because of the nature of the continuous time model (see, e.g., Chen [1992] and Zhang and Chen [1991]); (iii): the unknown parameters are always estimated no matter whether they are needed or not (see, e.g., Chen [1992]; Chen and Zhang [1992]; and Zhang and Chen [1991]); (iv) some external stochastic excitation signals are invoked (see, e.g., Chen and Zhang [1992] and Zhang and Chen [1992]).

In this paper, we formulate an adaptive control algorithm which (i) avoids use of $dx_t$ or $dy_t$ and the introduction of external stochastic signals in the procedures of parameter estimation and adaptive control, (ii) simplifies the criteria in steps A through C so that they are required to be verified at discrete time instants only, (iii) stop the parameter estimation procedure when it is not needed in order to make the adaptive control law more practical, and finally, (iv) does not use an external stochastic excitation signal. (In effect the Brownian motion $w$ driving the system is exploited for this purpose.)

It may be conjectured that such an alternation of identification and control regimes will work in certain time-varying cases.

2. Full observation systems. In this section we consider the LQ adaptive control problem for the following system model:

\[ dx_t = Ax_t dt + Bu_t dt + Cdwt, \quad t \geq 0, \]

where $x_t \in \mathbb{R}^n$ and $u_t \in \mathbb{R}$ are the state and input of the system and \{wt, \mathcal{F}_t\} is a standard Wiener process in \mathbb{R}^m.

Using controls which at any instant $t$ are based only on information available up to time $t$, we wish to stabilize the system (2.1). In this paper, this is achieved by the use of controls which are certain time-interleaved versions of an excitation signal and a signal designed via the certainty equivalence principle for the following quadratic loss function:

\[ \min_{u \in U} \lim_{t \to \infty} \sup_{\tau \leq t} J_t(u), \]

where

\[ U = \{ u_t : u_t \in \mathcal{F}_t \triangleq \sigma\{x_\tau, u_s, 0 \leq \tau \leq t, 0 \leq s < t\}, t \geq 0 \}, \]

\[ J_t(u) = \frac{1}{t} \int_0^t (x_s^T Q_1 x_s + Q_2 u_s^2) ds, \quad Q_1 \geq 0, \quad Q_2 > 0. \]

This problem has been investigated in previous work; see, for instance, Zhang and Chen [1991], Chen [1992], where these authors presented the first rigorous stability analysis for such adaptive stabilization of a system, which might be neither open-loop stable nor minimum phase and might be subject to disturbances with an unknown bound.

Specification of the adaptive control law. First, we define a causal system which shall generate a disturbance input $u'$, which shall be employed over an at most countable set of intervals; second, we define a linear state feedback control law, which shall be used over the intervals which interleave those during which the disturbance
is used; then, third, we give a rule for determining the switching times which depend solely upon the history of the system inputs and outputs.

Following the procedure described from step A to step D, we now find an adaptive stabilization control for model (2.1).

Assume that the control input $u$ is defined on the interval $[0, t)$; then the input $u'$ and a countable sequence of stopping times with no finite accumulation point are defined as follows:

Let $T > 1$ and $\alpha$ be positive constants chosen arbitrarily. Define for $i = 1, 2, \ldots, n+1$

$\beta_i = (-1)^{i+1} \alpha^i \frac{(n+1)!}{i! \times (n+1-i)!}$ \hspace{1cm} with \hspace{1cm} $0! \triangleq 1, \hspace{1cm} i! \triangleq 1 \times 2 \times \cdots \times i$

and for some integer $k \geq 0$

$u'_t = L_{T^k} + \beta_1 S_k u'_t + \cdots + \beta_{n+1} S_{n+1} u'_t, \hspace{1cm} t \in [T^k, T^{k+1})$

where $S_k u_t = \int_{T^k}^t u_s ds$ and $L_t = 1 + \int_0^t (\|\psi_s\|^2 + \|z_s\|^2 + u^2_s) ds$ with $\psi_t = [Sx^T_t, Su_t]^T$ and $z_t = [Sx^T_t, Su_t]^T$. Here, $S$ is the integral operator $Sx_t = \int_0^t x_s ds$; $x_t$ and $u_t$ are, respectively, the system state and system input, which is recursively given by (2.5)–(2.12); $\bar{x}_t$ and $\bar{u}_t$ are the solutions of the following equations, respectively:

$(S + 1)\bar{x}_t = x_t, \hspace{1cm} (S + 1)\bar{u}_t = u_t,$

i.e.,

$\bar{x}_t = x_t - \int_0^t e^{-(t-\lambda)} x_{\lambda} d\lambda, \hspace{1cm} \bar{u}_t = u_t - \int_0^t e^{-(t-\lambda)} u_{\lambda} d\lambda.$

It will be seen below that the function $u$ which appears in the definitions above is equal to $u'$ during the time intervals when the excitation input to the system is in use and is given as a linear function of the state $x$ during the periods when $u'$ is not being used as a system input.

Set $\theta = [A, B]^T$. Choosing an arbitrary $\theta_0$, the unknown parameter $\theta$ is estimated via the least-squares method, which is modified so as to be active only over a sequence of intervals $[T^{\tau_i-1}, T^{\tau_i})$. Specifically, the estimate $\hat{\theta}_t = [A_t, B_t]^T$ is given by

$\hat{\theta}_t = P_t \psi'_t (\bar{x}_t - \theta_0^T \psi'_t)$ \hspace{1cm} with \hspace{1cm} $P_t = \left(I + \int_0^t \psi'_s \psi''_s ds\right)^{-1}$

and

$\psi'_t = \begin{cases} z_t & \text{if } t \in [0, T^{\tau_0}) \text{ or } t \in [T^{\tau_{i-1}}, T^{\tau_i}) \\ 0 & \text{if } t \in [T^{\sigma_i}, T^{\tau_i}) \end{cases}$ \hspace{1cm} \forall i \geq 1,$

where $\{\tau_i\}$ and $\{\sigma_i\}$ are two stopping time sequences defined as follows: $0 = \tau_0 < \sigma_1 < \tau_1 < \sigma_2 < \tau_2 < \cdots$.

$\sigma_i = \inf\{k > \tau_{i-1} : k \in \mathcal{N}, \hspace{0.5cm} (A_{T^k}, B_{T^k}, D) \text{ is controllable and observable} \}$

where here, and hereafter, $\mathcal{N}$ denotes the set of all positive integers, and $D$ is any square matrix such that $D^T D = Q_1$. 

(2.9)
\[ \tau_i = \inf \left\{ k > \sigma_i : k \in \mathcal{N}, \quad \int_0^{T_k} \|x_s\|^2 ds > T^{\sigma_i} \int_0^{T^{\sigma_i}} (\|x_s\|^2 + \|u''_s\|^2) ds + T^{\sigma_i} T^k + T^{\sigma_i} \right\} \]

(2.10)

with the excitation input \(u''\) given by

\[ u''_t = \begin{cases} u'_t, & \text{if } t \in [0, T^{\tau_0}) \text{ or } t \in [T^{\tau_{i-1}}, T^{\sigma_i}) \\
0, & \text{if } t \in [T^{\sigma_i}, T^{\tau_i}) \end{cases} \quad \forall i \geq 1. \]

(2.11)

Here \(Q_1\) is given in (2.4) and \(u'_t\) is defined by (2.6).

The adaptive control \(u\) is generated by interleaving the excitation input \(u'\) and a linear feedback input as follows:

\[ u_t = \begin{cases} u'_t, & \text{if } t \in [0, T^{\tau_0}) \text{ or } t \in [T^{\tau_{i-1}}, T^{\sigma_i}) \text{ for some } i \geq 1, \\
-Q_2^{-1} B_{T^{\sigma_i}} R_{T^{\sigma_i}} x_t, & \text{if } t \in [T^{\sigma_i}, T^{\tau_i}) \text{ for some } i \geq 1,
\end{cases} \]

(2.12)

where \(Q_2\) is the positive constant in (2.4), \(B_{T^{\sigma_i}}\) is the estimate for \(B\) at time instant \(T^{\sigma_i}\) given by (2.7) and (2.8), and \(R_{T^{\sigma_i}}\) is a solution of the following algebraic Riccati equation:

\[ A_{T^{\sigma_i}} R_{T^{\sigma_i}} + R_{T^{\sigma_i}} A_{T^{\sigma_i}} - R_{T^{\sigma_i}} B_{T^{\sigma_i}} Q_2^{-1} B_{T^{\sigma_i}}^T R_{T^{\sigma_i}} + D^T D = 0. \]

(2.12)

Here \(A_{T^{\sigma_i}}\) is the estimate for \(A\) at time instant \(T^{\sigma_i}\) given by (2.7) and (2.8).

Remark 2.1. From the definition (2.12) of \(u_t\), it is easy to see that \(\tau_i\) and \(\sigma_i\) are Markov times, i.e., \(\sigma\{T^{\tau_i} \leq t\} \in \mathcal{F}_t\) and \(\sigma\{T^{\tau_i} \leq t\} \in \mathcal{F}_t\). Thus, \(u \in \mathcal{U}\) and \(\psi'_t \in \mathcal{F}_t\).

Remark 2.2. The excitation signal in (2.12) is generated by (2.6), in which a designer needs only determine the deterministic coefficients \(\beta_i\). No additional stochastic signal is introduced except \(L_{T^k}\), so we call this a deterministic-like excitation signal.

Remark 2.3. In (2.9), to get \(\sigma_i\), the only thing one should do is to check the controllability and observability of \((A_{T^k}, B_{T^k}, D)\) for every integer \(k\) at time instant \(T^k\). While in (2.10), one need only check whether or not

\[ \int_0^{T_k} \|x_s\|^2 ds > T^{\sigma_i} \int_0^{T^{\sigma_i}} (\|x_s\|^2 + \|u''_s\|^2) ds + T^{\sigma_i} T^k + T^{\sigma_i} \]

for every integer \(k\) at time instant \(T^k\) and such a set of time instants is evidently countable.

Remark 2.4. By (2.7) and (2.8) it is easy to see that \(\theta_t = \theta_{T^{\sigma_i}}\) for all \(t \in [T^{\sigma_i}, T^{\tau_i})\). In other words, unlike in Zhang and Chen [1991] or in Chen [1992], the LS parameter estimation is not carried out in the time interval \([T^{\sigma_i}, T^{\tau_i})\). Thus, if adaptive control (2.12) results in an integer \(i\) such that \(\sigma_i < \infty\) and \(\tau_i = \infty\), then the unknown parameter estimates will be locked on an acceptable value \(\theta_{T^{\sigma_i}}\) forever.

Lemma 2.1. Let \(\lambda_{\min}^{(t)}\) denote the smallest eigenvalue of matrix \(P_{(t)}^{-1}\). Then the parameter estimate \(\theta_t\) given by (2.7)–(2.12) has the following property:

\[ \|\theta_t - \theta\|^2 \leq \frac{c(t+1)}{\lambda_{\min}^{(t)}} \quad \forall t \geq 0, \]

where here and hereafter \(c \geq 0\) is a possibly random quantity which is independent of \(t\).
Proof. Let $\tilde{\theta}_t = \theta_t - \theta$. Then from (2.1) it follows that
\[
x_t = x_0 + \theta^\top \psi_t + Cw_t.
\]
Substituting this into the first equation of (2.7) and noting (2.8), $(S + 1)\bar{x}_t = x_t$, and $(S + 1)z_t = \psi_t$ we get
\[
\dot{\tilde{\theta}}_t = -P_t \psi_t^\top \psi_t^\top \tilde{\theta}_t + P_t \psi_t^\top [(S + 1)^{-1}(Cw_t) + \varepsilon_t],
\]
where here and hereafter $\varepsilon_t$ denotes a time function which exponentially converges to zero.

From this and the second equation of (2.7) we obtain
\[
\frac{d(\tilde{\theta}_t^\top P_t^{-1} \tilde{\theta}_t)}{dt} = - (\tilde{\theta}_t^\top \psi_t^\top)^2 + 2 \tilde{\theta}_t^\top \psi_t^\top [(S + 1)^{-1}(Cw_t) + \varepsilon_t],
\]
which implies that
\[
0 \leq \tilde{\theta}_t^\top P_t^{-1} \tilde{\theta}_t = \tilde{\theta}_0^\top P_0^{-1} \tilde{\theta}_0 - \int_0^t (\tilde{\theta}_s^\top \psi_s^\top)^2 ds + 2 \int_0^t \tilde{\theta}_s^\top \psi_s^\top [(S + 1)^{-1}(Cw_s) + \varepsilon_s] ds
\]
where for the last inequality we have invoked $\int_0^t [(S + 1)^{-1}(Cw_s)]^2 ds = O(t + 1)$ a.s.
(e.g., Chen and Guo [1990]). Therefore, Lemma 2.1 is true.

Lemma 2.2. In system (2.1), if $(A, B)$ is controllable and $u_t = u_t^*$ for some $k$ and all $t \in (T^k, T^{k+1}]$, then there exist $c > 0$, $a > 1$, and $k_0 > 0$ such that
\[
\lambda_{\text{min}}^{(T^{k+1})} \geq ca^{T^{k+1}} L_{T^k} \quad \forall k \geq k_0.
\]

Proof. See Appendix A.

Theorem 2.1. If $\text{Span}(B) \subset \text{Span}(C)$ and $(A, B, D)$ is controllable and observable with $D^\top D = Q_1$, then under the adaptive control (2.12) it is the case that
(a) there is an integer $i$ such that $\tau_i < \infty$, $\tau_i = \infty$, and $\theta_t = \theta_{T^{\sigma_i}}$, $\forall t \geq T^{\sigma_i}$;
(b) $\lim_{t \to \infty} \frac{1}{t} \int_0^t x_s^\top x_s^\top ds$ and $\lim_{t \to \infty} \frac{1}{t} \int_0^t u_s^\top u_s^\top ds$ exist and are finite a.s.
Here and hereafter by $\text{Span}(X)$ we mean the linear space spanned by the column vectors of $X$.

Proof. Let $R$ be the solution of the following algebraic Riccati equation:
\[
A^\top R + RA - RBQ_2^{-1}B^\top R + D^\top D = 0.
\]
Then it is well known that $\Phi = A - BQ_2^{-1}B^\top R$ is stable, and hence there is a positive matrix $P$ such that
\[
P\Phi + \Phi^\top P = -I.
\]
From this, it is easy to see that there exists a small enough positive constant $\varepsilon$ such that
\[
P\Phi_t + \Phi_t^\top P \leq -\frac{1}{2} I \quad \forall \Phi_t \in \{\Phi_t : \|\Phi_t - \Phi\| < \varepsilon\}.
\]
We now define

\[ \Phi_t = \begin{cases} 
A & \text{if } t \in [0, T^\tau) \text{ or } t \in [T^\tau_i - 1, T^\tau_i), \\
A - BQ_{2^{-1}}R_{T^\tau_i} & \text{if } t \in [T^\tau_i, T^\tau_i). 
\end{cases} \]

Then the adaptive control system defined above is expressed as

\[ dx_t = \Phi_t x_t dt + Bu''_t dt + Cdw_t. \]

From the definitions (2.9) and (2.10) of sequences \( \{\sigma_i\} \) and \( \{\tau_i\} \) we see that only three cases can possibly hold. The first case is that there exists an integer \( i \) such that \( \tau_{i-1} < \infty \) and \( \sigma_i = \infty \); the second is that \( \sigma_i < \tau_i < \infty \) for every integer \( i \geq 1 \), and the third is there exists an integer \( i \) such that \( \sigma_i < \infty \) and \( \tau_i = \infty \). We shall now show the first two cases are impossible.

**Case 1.** It is impossible that an integer \( i \) such that \( \tau_{i-1} < \infty \) and \( \sigma_i = \infty \) exists.

From Lemmas 2.1 and 2.2, it is easy to see that if there were an integer \( i \) such that \( \tau_{i-1} < \infty \) and \( \sigma_i = \infty \), then there would be

\[ A_{T^k} \to A \text{ and } B_{T^k} \to B. \]

Thus, by the assumption that \( (A, B, D) \) is controllable and observable, we see that there would exist a \( k \) such that \( T^k \geq T^\tau_i - 1 \) and \( (A_{T^k}, B_{T^k}, D) \) is controllable and observable. This contradicts \( \sigma_i = \infty \).

**Case 2.** It is impossible that \( \sigma_i < \tau_i < \infty \) for every integer \( i \geq 1 \).

If for every integer \( i \geq 1 \), \( \sigma_i < \tau_i < \infty \), then by Lemmas 2.1 and 2.2 it is easy to see that \( \Phi_{T^\tau_i} \to \Phi \). Therefore, there exists \( i_0 \) such that for all \( i \geq i_0 \),

\[ \|\Phi_{T^\tau_i} - \Phi\| < \varepsilon, \]

which together with (2.14) implies that

\[ \|\Phi_t\| \leq c \text{ and } P\Phi_{T^\tau_i} + \Phi_{T^\tau_i}^T P \leq -\frac{1}{2} I. \]

Using Ito’s formula (cf. Schwartz [1984]) we find that for \( k \in [\sigma_i, \tau_i] \cap N \)

\[ x^T_{T^k} P x_{T^k} \leq x^T_0 P x_0 + \int_0^{T^\tau_i} x^T_s (P \Phi_s + \Phi^T_s P) x_s ds - \frac{1}{2} \int_{T^\tau_i}^{T^k} \|x_s\|^2 ds \]

\[ + 2 \int_0^{T^\tau_i} x^T_s P B u''_s ds + 2 \int_0^{T^k} x^T_s P C dw_s + \text{tr}(C^T P C) T^k, \]

where here and hereafter \( \text{tr}(X) \) denotes the trace of \( X \).

Note that by Lemma 4 of Christopeit [1986], there exist random numbers \( c', c'' \), independent of \( t \), such that for all \( k \) sufficiently large, say, for all \( k \geq \sigma_m \),

\[ \int_0^{T^k} x^T_s P C dw_s \leq c' \left( \int_0^{T^k} \|x_s\|^2 ds \right)^{\eta + \frac{1}{2}} + c' \leq c'' + \frac{1}{8} \int_0^{T^k} \|x_s\|^2 ds, \quad \eta \in \left( 0, \frac{1}{2} \right). \]

Then from (2.16), for some random number \( c''' \) independent of time, we have

\[ x^T_{T^k} P x_{T^k} \leq c''' + c'' \int_0^{T^\tau_m} \|x_s\|^2 ds - \frac{3}{8} \int_{T^\tau_m}^{T^k} \|x_s\|^2 ds \]

\[ + c \int_0^{T^\tau_m} \|u'_s\|^2 ds + \text{tr}(C^T P C) T^k, \]
and hence, for some random number \( c \) independent of time,

\[
\int_0^{T_k} \|x_s\|^2 ds \leq c \int_0^{T_m} \left( \|x_s\|^2 + \|u'_s\|^2 \right) ds + cT_k + c \quad \forall k \geq \sigma_m.
\]

Now there exists \( i \) sufficiently large that \( T^{\sigma_i} \leq c \), and so by (2.17) this gives

\[
\int_0^{T_k} \|x_s\|^2 ds \leq T^{\sigma_i} \int_0^{T^{\sigma_i}} \left( \|x_s\|^2 + \|u'_s\|^2 \right) ds + T^{\sigma_i}T_k + T^{\sigma_i} \quad \forall k \geq \sigma_i,
\]

which contradicts \( \tau_i < \infty \).

So, there must exist an integer \( i \) such that \( \sigma_i < \infty \) and \( \tau_i = \infty \).

Thus by (2.7) and (2.8) we get Assertion (a) of Theorem 2.1. Furthermore, (2.10) together with \( \tau_i = \infty \) implies that

\[
\limsup_{k \to \infty} \frac{1}{T_k} \int_0^{T_k} \|x_s\|^2 ds < \infty \quad \text{a.s.}
\]

Notice that \( \text{Span}(B) \subseteq \text{Span}(C) \) and \((A, B)\) is controllable implies that \((\Phi_{T^{\sigma_i}}, C)\) is controllable. Then by (2.18) and Lemma B.1 in Appendix B we see that \( \Phi_{T^{\sigma_i}} \) is stable. Therefore, from Lemma 3 of Chen and Guo [1990] it follows that

\[
\lim_{t \to \infty} \frac{1}{t} \int_0^t x_s x_s^T ds \quad \text{exists and is finite a.s.,}
\]

which together with (2.12) implies that

\[
\lim_{t \to \infty} \frac{1}{t} \int_0^t u_s^2 ds = Q_S^{-2} B_{T^{\sigma_i}}^T R_{T^{\sigma_i}} \left( \lim_{t \to \infty} \frac{1}{t} \int_0^t x_s x_s^T ds \right) (B_{T^{\sigma_i}}^T R_{T^{\sigma_i}})^T \quad \text{exists and is finite a.s.}
\]

This proves Assertion (b) of Theorem 2.1. \( \square \)

3. Partially observed system.

3.1. Problem statement. In this section we consider the single-input single-output continuous-time system described by

\[
A(S)y_t = y_0 + SB(S)u_t + C(S)w_t + S\eta_t \quad \forall t \geq 0,
\]

where \( A(S), B(S), \) and \( C(S) \) are polynomials in \( S \) with unknown coefficients:

\[
A(S) = 1 + \sum_{i=1}^p a_i S^i, \quad B(S) = \sum_{i=1}^q b_i S^{i-1}, \quad C(S) = \sum_{i=0}^l c_i S^i;
\]

\( \{w_t, \mathcal{F}_t\} \) is a standard Wiener process with respect to a nondecreasing \( \sigma \)-algebras \( \{\mathcal{F}_t\} \) defined on a probability space; \( y_t \) and \( u_t \) are the system output and input, respectively, and measurable with respect to \( \mathcal{F}_t \); and \( \eta_t \) is unknown disturbance or unmodeled dynamics which is measurable with respect to \( \mathcal{F}_t \).

As Zhang and Chen (1991) show, model (3.1) subject to (3.2) is very general and includes some widely used models. For instance, in the case where \( l = p \) and \( c_p = ga_p \), the input-output properties of (3.1) are equivalent to those of the following well-known...
state-space representation (e.g., Gevers, Goodwin, and Wertz [1991]; Goodwin, et al. [1991]; Caines [1992]):

\[ dx_t = Ax_t dt + Bu_t dt + Cd w_t + D \eta_t dt, \]
\[ dy_t = D^* x_t dt + gw_t, \]

with

\[
A = \begin{bmatrix}
-a_1 & 1 \\
-a_2 & & \\
n & & & \\
-a_m & & & 1
\end{bmatrix}, \quad
B = \begin{bmatrix}
b_1 \\
b_2 \\
m & & \\
b_m
\end{bmatrix}, \quad
C = \begin{bmatrix}
c_0 - go_0 \\
& & \\
& & \\
& & c_{m-1} - ga_{m-1}
\end{bmatrix}, \quad
D = \begin{bmatrix}
1 \\
0 \\
& & \\
& & 0
\end{bmatrix},
\]

where \( m = \max\{p, q\} \) and here and hereafter we set \( a_0 = 1, a_i = 0 \) for \( i > p \), \( b_j = 0 \) for \( j > q \), and \( c_k = 0 \) for \( k > l \).

Let us denote the collection of unknown coefficients of \( A(S) \) and \( B(S) \) by \( \theta \):

\[
\theta = [-a_1, \ldots, -a_p, b_1, \ldots, b_q]^\top.
\]

Let

\[ F(S) = 1 + f_1 S + \cdots + f_{l+1} S^{l+1} \]

be an arbitrarily given stable polynomial of \( S \); i.e., every \( S \) that satisfies \( F(S) = 0 \) has negative real part.

Denote by \( y^f_t \) and \( u^f_t \) the filtered value, respectively:

\[ F(S)y^f_t = y_t, \quad F(S)u^f_t = u_t \]

and

\[ \varphi^f_t = [S y^f_t, \ldots, S^p y^f_t, \quad S u^f_t, \ldots, S^q u^f_t]^\top. \]

Define

\[
\varphi^f_t = \begin{cases}
\varphi^f_t & \text{if } t \in [0, T_{\tau_0}) \text{ or } t \in [T_{\tau_{i-1}}, T_{\tau_i}) \text{ for some } i \geq 1, \\
0 & \text{if } t \in [T_{\tau_i}, T_{\tau_{i+1}}) \text{ for some } i \geq 1,
\end{cases}
\]

where \( \{\tau_i\} \) and \( \{\sigma_i\} \) are two stopping time sequences such that \( \varphi^f_t \in \mathcal{F}_t \).

Then the unknown parameter \( \theta \) is estimated as follows:

\[ \hat{\theta}_t = P_t \varphi^f_t (y^f_t - \theta_t \varphi^f_t) \quad \text{with} \quad P_t = \left( I + \int_0^t \varphi_s \varphi_s^\top ds \right)^{-1}, \]

where \( \theta_0 \) is a constant chosen arbitrarily.

The purpose of this paper is to design a \( \theta_t \)-based adaptive control so that the closed-loop system is stabilized in the sense that

\[ \sup_{t \geq 0} \frac{1}{t+1} \int_0^t (y^2_s + u^2_s) ds < \infty \quad \text{a.s.} \]

under the following assumptions:
A.1. \( A(S) \) and \( SB(S) \) are coprime, \( b_q \neq 0, \ l \leq \min\{p, q - 1\} \), and \( p \) and \( q \) are known.

A.2. \( \sup_{t \geq 0} \int_{t+1}^{t} \eta_2^2 ds < \infty \).

From Zhang and Chen [1991] it follows that Assumptions A.1 and A.2 are as weak as the following necessary and sufficient ones even when \( \theta \) is known:

A.1'. The greatest common factor of \( A(S) \) and \( SB(S) \) is 1 or a stable polynomial, \( b_q \neq 0, \) the order of the greatest common factor, \( l \leq \min\{p, q - 1\} \), and \( p \) and \( q \) are known.

A.2'. \( \sup_{t \geq 0} \int_{t+1}^{t} \eta_2^2 ds < \infty \) a.s.

However, for simplicity of notation, in this paper we use Assumptions A.1 and A.2.

Remark 3.1. We now look at how to calculate the filtered values \( y^f_t \) and \( u^f_t \) of \( y_t \) and \( u_t \), respectively, with respect to filter \( F(S) \) in (3.4).

Let

\[
D_F = \begin{bmatrix}
-f_1 & -f_2 & \cdots & -f_{t+1} \\
1 & 0 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots \\
0 & 1 & 0
\end{bmatrix}, \quad Y_t = \begin{bmatrix}
y^f_t \\
S^t y^f_t \\
\vdots \\
S^t y^f_t
\end{bmatrix}, \quad H_t = \begin{bmatrix}
1 \\
0 \\
\vdots \\
0
\end{bmatrix}
\]

Then from (3.4) we see that

\[
Y_t = D_F S Y_t + H_t y_t,
\]

which is equivalent to

\[
Y_t = D_F \int_0^t e^{D_F(t-\lambda)} H_t y_t d\lambda + H_t y_t.
\]

Thus we have

\[
y^f_t = H_t^T Y_t = y_t + H_t^T D_F \int_0^t e^{D_F(t-\lambda)} H_t y_t d\lambda.
\]

Similarly, we can get

\[
 u^f_t = u_t + H_t^T D_F \int_0^t e^{D_F(t-\lambda)} H_t u_t d\lambda.
\]

3.2. Adaptive control. We first look at what the stabilization control is in the case where \( \theta \) is known. To see this, we introduce the following lemma (e.g., Chen and Guo [1990]).

Lemma 3.1. Let \( k \geq 0 \) be an integer and \( E(S) = 1 + e_1 S + \cdots + e_k S^k \) be a stable polynomial with \( e_k \neq 0 \). Then there is a (nonrandom) constant \( C_e \geq 1 \) (depending on \( E(S) \) only) such that

\[
\sum_{i=0}^{k} \int_0^t \left( \frac{S^i}{E(S)} z_\lambda \right)^2 d\lambda \leq C_e \int_0^t z^2_\lambda d\lambda
\]

for any square-integrable process \( \{z_\lambda\} \).

If \( A(S), SB(S) \) are coprime and \( b_q \neq 0, \) then for any polynomial

\[
E(S) = 1 + e_1 S + \cdots + e_{p+q} S^{p+q} \text{ with } e_{p+q} \neq 0,
\]
there exists a unique polynomial pair \((G(S), H(S))\) such that

\[
(3.10) \quad A(S)G(S) - SB(S)H(S) = E(S) \quad \text{with } \partial(G(S)) \leq q - 1 \text{ and } \partial(H(S)) = p,
\]
where here and hereafter \(\partial(X(S))\) denotes the degree of the polynomial \(X(S)\) in \(S\).

From (3.10) and (3.1) it is clear that

\[
E(S)y_t = A(S)G(S)y_t - SB(S)H(S)y_t
\]

\[
= G(S)[A(S)y_t - SB(S)u_t] + SB(S)[G(S)u_t - H(S)y_t]
\]

\[
= G(S)[y_0 + C(S)w_t + S\eta_t] + SB(S)[G(S)u_t - H(S)y_t]
\]

and

\[
E(S)u_t = A(S)G(S)u_t - SB(S)H(S)u_t
\]

\[
= H(S)[A(S)y_t - SB(S)u_t] + A(S)[G(S)u_t - H(S)y_t]
\]

\[
= H(S)[y_0 + C(S)w_t + S\eta_t] + A(S)[G(S)u_t - H(S)y_t].
\]

Therefore, in the case where \(\theta\) is known, \(l \leq \min\{p, q - 1\}\) and Assumptions A.1 and A.2 hold, for any given stable \(E(S)\) subject to (3.9), if the control is defined as follows:

\[
G(S)u_t - H(S)y_t = 0, \quad t \geq 0,
\]
then the system is stabilized in the average sense (3.8).

Similar to §2, we now introduce a deterministic-like excitation signal \(u'_t\). Let \(T > 1\) and \(\alpha\) be positive constants chosen arbitrarily. Define for \(i = 1, 2, \ldots, p + q\),

\[
\beta_i = (-1)^{i+1}\alpha^i \frac{(p + q)!}{i! \times (p + q - i)!} \quad \text{with } 0! \triangleq 1, \quad i! \triangleq 1 \times 2 \times \cdots \times i,
\]
and for some integer \(k \geq 0\),

\[
u'_t = L_{T^k} + \beta_1 S_k u'_t + \cdots + \beta_{p+q} S_k^{p+q} u'_t, \quad t \in [T^k, T^{k+1}),
\]
where \(S_k u_t = \int_{T^k}^{T^k} u_s ds\) and \(L_t = 1 + \int_0^t (\|\varphi_s^t\|^2 + \|\varphi_s\|^2 + u_s^2) ds\) with \(\varphi_t^f\) defined by (3.5) and \(\varphi_t\) defined by

\[
(3.11) \quad \varphi_t = [S_{yt}, \ldots, S_{p_yt}, Su_t, \ldots, S_{q_u}u_t]'.
\]

For any \(k \geq 1\), write \(\theta_{T^k}\) in the component form

\[
\theta_{T^k} = [-a_{1T^k}, \ldots, -a_{pT^k}, b_{1T^k}, \ldots, b_{qT^k}]'
\]
and set

\[
(3.12) \quad A_k(S) = 1 + \sum_{i=1}^p a_{iT^k} S^i, \quad B_k(S) = \sum_{i=1}^q b_{iT^k} S^{i-1}.
\]

Let \(E(S)\) be a stable polynomial subject to (3.9). Then by Lemma 3.1 there is a constant \(C_e\) (depending on \(E(S)\) only) such that

\[
(3.13) \quad \sum_{i=0}^{p+q} \int_0^t \left( \frac{S^i}{E(S)} \right)^2 \frac{d\lambda}{E(S)} \leq C_e \int_0^t \frac{z^2}{\lambda} d\lambda.
\]
Let $R(S)$ be a stable polynomial of $S$ with $\partial (R(S)) = \min \{p + 1, q\}$ and let $\xi^u_t$, $\xi^y_t$ denote the filtered values of $u_t$, $y_t$ with respect to $R(S)$, respectively:

$$R(S)\xi^u_t = u_t, \quad R(S)\xi^y_t = y_t \quad \forall t \geq 0.$$  

(3.14)

Actually, the filtered values $\xi^u_t$ and $\xi^y_t$ can be calculated as in Remark 3.1.

Set $\zeta_t = [S\xi^y_t, \ldots, S^p\xi^y_t, S^{p+1}\xi^y_t, \ldots, S^q\xi^y_t]^T$. Then by (3.11) and (3.14) we get

$$R(S)\zeta_t = \varphi_t.$$  

(3.15)

In the following discussion, for a given polynomial $Z(S) = z_0 + z_1 S + \cdots + z_r S^r$, its norm is defined as

$$\|Z(S)\| = \left( \sum_{i=0}^r z_i^2 \right)^{1/2}.$$  

Define switching times $1 = \tau_0 < \sigma_1 < \tau_1 < \sigma_2 < \tau_2 \cdots$ as follows:

$$\sigma_i = \inf \left\{ k > \tau_{i-1} : \begin{array}{l} A_k(S)G_k(S) - SB_k(S)H_k(S) = E(S) \text{ is solvable} \\ \partial(G_k(S)) \leq q - 1 \quad \text{and} \quad \partial(H_k(S)) = p; \quad \text{and} \\ \|G_k(S)\|^2 + \|H_k(S)\|^2 \leq \frac{k}{2C_c(p + q + 1)}; \end{array} \right\}$$  

(3.16)

$$\tau_i = \min \left\{ k > \sigma_i : \begin{array}{l} \text{there exists} \ t \in (T_{\tau_i}, T^k) \text{ such that} \\ \int_0^{T^k} [\xi^y_s - \theta^*_{T,T^k}\zeta_s]^2 ds \leq k^{-2}f(k, T^k) \end{array} \right\},$$  

(3.17)

where $C_c$ is given in (3.13), and

$$f(x, t) = (t + 1) \sup_{0 \leq \lambda \leq t} \left\{ x^3 + \frac{1}{\lambda + 1} \int_0^{p+1} \sum_{j=0}^{\lambda} (S^j \xi^y_s)^2 + \sum_{j=0}^{q} (S^j \xi^u_s)^2 \right\} ds.$$  

(3.18)

Similar to (2.12) we define the adaptive control $u_t$ as follows:

$$u_t = \begin{cases} \begin{array}{ll} u'_t & \text{if} \ t \in [0, T^{\tau_0}) \text{ or } t \in [T^{-\tau_{i-1}}, T^{\sigma_i}) \text{ for some } i \geq 1, \\
H_{\sigma_i}(S)y_t - (G_{\sigma_i}(S) - 1)u_t & \text{if} \ t \in [T^{\sigma_i}, T^\tau_i) \text{ for some } i \geq 1, \end{array} \end{cases}$$  

(3.19)

where $H_{\sigma_i}(S)$ and $G_{\sigma_i}(S)$ are generated by (3.16), (3.12), and (3.4)–(3.7).

In this case, similar to Lemmas 2.1 and 2.2 we may obtain the following results.

**Lemma 3.2.** Let $\lambda_{\min}^{(t)}$ denote the smallest eigenvalue of matrix $R_t^{-1}$. Then the parameter estimate $\theta_t$ given by (3.4)–(3.7) has the following property:

$$\|\theta_t - \theta\|^2 \leq \frac{c(t + 1)}{\lambda_{\min}^{(t)}} \quad \forall t \geq 0.$$  

(3.20)
for some time-independent random variable \( c \geq 0 \).

Proof. Let \( \tilde{\theta}_t = \theta_t - \theta \). Then from (3.1)–(3.3) and (3.5) it follows that

\[
y_t^f = \theta^T \varphi_t^f + F^{-1}(S)[C(S)w_t + S\eta_t + \varepsilon_t],
\]

where we recall that \( \varepsilon_t \) denotes a function of time decaying exponentially to zero. Substituting this into the first equation of (3.7) and noting (3.6), we get

\[
\dot{\tilde{\theta}}_t = -P_t \tilde{\varphi}_t \tilde{\varphi}_t^T \tilde{\theta}_t + P_t \tilde{\varphi}_t \left[ F^{-1}(S)(C(S)w_t + S\eta_t) + \varepsilon_t \right].
\]

Then combining this with the second equation of (3.7) gives

\[
\frac{d(\tilde{\theta}_t^T P_t^{-1} \tilde{\theta}_t)}{dt} = - (\tilde{\theta}_t^T \tilde{\varphi}_t)^2 + 2\tilde{\theta}_t^T \tilde{\varphi}_t \left[ F^{-1}(S)(C(S)w_t + S\eta_t) + \varepsilon_t \right],
\]

which together with \( \int_0^t [F^{-1}(S)(C(S)w_s)]^2 ds = O(t) \) (cf. Lemma 3 of Chen and Guo [1990]) implies that

\[
\frac{1}{c} \int_0^t \left| \tilde{\theta}_t^T \tilde{\varphi}_t \right|^2 ds \leq O(t + 1),
\]

where for the final bound we have used Lemma 3.1 and Assumption A.2.

**Lemma 3.3.** Under Assumptions A.1 and A.2, if \( u_t = u_t' \) for some \( k \) and any \( t \in (T^k, T^{k+1}] \), then there exist \( c > 0 \), \( a > 1 \), and \( k_0 > 0 \) such that

\[
\lambda_{\text{min}}^{(T^{k+1})} \geq caT^{k+1} L_{T^k} \quad \forall k \geq k_0.
\]

**Proof.** The proof resembles that of Lemma 2.2 and is given in Appendix C.

**Theorem 3.1.** Under Assumptions A.1 and A.2 and the adaptive control (3.4)–(3.7), (3.16)–(3.19), we get that

(a) there is an integer \( k \) such that \( \sigma_i < \infty \), \( \tau_i = \infty \), and \( \theta_t = \theta_{T^i} \), \( \forall t \geq T^i \); and

(b) \( \sup_{t \geq 0} \frac{1}{T^{k+1}} \int_0^t (y_t^2 + u_t^2) ds < \infty \) a.s.

**Proof.** We first show that it is impossible that \( \tau_i < \infty \) and \( \sigma_{i+1} = \infty \) on a sample set \( \mathcal{D} \) with positive probability for an integer-valued random variable \( \tau \geq 0 \).

In fact, if there were a sample set \( \mathcal{D} \) of positive probability, i.e., for which \( P(\mathcal{D}) > 0 \), which was such that for every sample \( \omega \in \mathcal{D} \), there were an \( i(\omega) \geq 0 \) (for simplicity, we drop \( \omega \) below) such that \( \tau_i < \infty \) and \( \sigma_{i+1} = \infty \), then \( u_t = u_t' \) for all \( t \geq \tau_i \). Thus, by Lemmas 3.2 and 3.3 we would have that for some constant \( a > 1 \)

\[
\|\theta_{T^k} - \theta\|^2 = O \left( \frac{T^k}{a^{T^k} L_{T^k}} \right) = O \left( \frac{1}{k^3} \right) \quad \text{a.s. on } \mathcal{D} \quad \forall k > \tau_i,
\]

which together with Lemma D.1 in Appendix D implies that there exists an integer \( k_1 \geq 0 \) such that for any \( k \geq k_1 \), \( A_k(S)G_k(S) - SB_k(S)H_k(S) = E(S) \) is solvable with respect to \( G_k(S) \) and \( H_k(S) \) subject to \( \partial(G_k(S)) \leq q - 1 \) and \( \partial(H_k(S)) = p \), and \( \|G_k(S)\|^2 + \|H_k(S)\|^2 \leq k/(2C_e(p + q + 1)) \).
From Lemma 3 of Chen and Guo [1990] and the fact that $\partial(R(S)) \geq \partial(C(S)) + 1$ it follows that

$$\int_0^t \left( \frac{C(S)}{R(S)} w_s \right)^2 ds = O(t) \quad \text{a.s.,}$$

while from (3.1), (3.3), (3.14), and (3.15) it follows that

$$\xi_s^y - \theta^r_{T^k} s = (\theta - \theta_t)^r s + \frac{C(S)}{R(S)} w_s + \frac{S}{R(S)} \eta_s + \varepsilon_t \quad \forall t, s \geq 0.$$

Therefore, by Assumption A.2, Lemma 3.1, and (3.18) we find that

$$\frac{1}{f(k, T^k)} \int_0^{T^k} \left[ \xi_s^y - \theta^r_{T^k} s \right]^2 ds \leq \frac{4}{f(k, T^k)} \left[ \int_0^{T^k} ((\theta - \theta_{T^k})^r s)^2 ds + \int_0^{T^k} \left( \frac{C(S)}{R(S)} w_s \right)^2 ds + \int_0^{T^k} \left( \frac{S}{R(S)} \eta_s \right)^2 ds + \int_0^{T^k} (\varepsilon_s)^2 ds \right]$$

$$= O \left( \|\theta_{T^k} - \theta\|^2 + \frac{1}{k^3} \right) = O \left( \frac{1}{k^3} \right) \quad \text{a.s. on } D,$$

where (3.20) is invoked for the last equality.

From (3.21) we conclude that there exists an integer $k_2 \geq k_1$ such that for any $k \geq k_2$,

$$\int_0^{T^k} \left[ \xi_s^y - \theta^r_{T^k} s \right]^2 ds \leq k^{-2} \quad \text{a.s. on } D.$$

Thus, $\sigma_{i+1} < \infty$ a.s. on $D$. This contradicts $\sigma_{i+1} = \infty$ on $D$ and $P(D) > 0$.

We now prove that $\tau_i = \infty$ a.s. for some integer-valued random variable $i \geq 1$.

In fact, from Lemmas 3.2 and 3.3 it follows that for some $a > 1$,

$$\|\theta_{T^\sigma_i} - \theta\|^2 = O \left( \frac{T^\sigma_i}{a^T^\sigma_i L_{T^\sigma_i}} \right) = O \left( \frac{1}{\sigma_i^3} \right).$$

As in (3.21) we would have

$$\int_0^t \left[ \xi_s^y - \theta^r_{T^\sigma_i} s \right]^2 ds = O \left( \frac{t^\sigma_i}{a^T^\sigma_i L_{T^\sigma_i}} \right) = O \left( \frac{1}{\sigma_i^3} \right) \leq \sigma_i^{-2},$$

where the last inequality is valid for some large enough $i$ and $t \geq T^\sigma_i$ because of (3.23). Hence there must be $\tau_i = \infty$ for some $i$; i.e., assertion (a) is true. We now prove assertion (b). From assertion (a) and (3.19) it follows that for some $i \geq 1$,

$$H_{\sigma_i}(S) y_t - G_{\sigma_i}(S) u_t = 0, \quad t \geq T^\sigma_i.$$

Henceforth, for simplicity of notation, we shall write $\theta^r_{\sigma_i}$ for $\theta^r_{T^k}$. 
In view of (3.16) we get

\[ E(S)S^k y_t = S^k A_\sigma(S)G_\sigma(S) y_t - S^{k+1} B_\sigma(S) H_\sigma(S) y_t \]
\[ = S^k G_\sigma(S)[A_\sigma(S) y_t - S B_\sigma(S) u_t] \]
\[ + S^{k+1} B_\sigma(S)[G_\sigma(S) u_t - H_\sigma(S) y_t], \quad k = 0, 1, \ldots, p + 1; \]
\[ E(S)S^k u_t = S^k H_\sigma(S)[A_\sigma(S) y_t - S B_\sigma(S) u_t] \]
\[ + S^{k+1} A_\sigma(S)[G_\sigma(S) u_t - H_\sigma(S) y_t], \quad k = 0, 1, \ldots, q. \]

Thus, noting that \( A_\sigma(S) y_t - S B_\sigma(S) u_t = y_t - \theta_\sigma, \xi_t, \) by (3.14), (3.25), and inequality \((a+b)^2 \leq 2a^2 + 2b^2\) we get

\[ S^j \xi_t^j = R(S)^{-1} S^j y_t \]
\[ = E(S)^{-1} S^j G_\sigma(S) R(S)^{-1}(y_t - \theta_\sigma, \xi_t) \]
\[ + E(S)^{-1} R(S)^{-1} S^{j+1} B_\sigma(S)[G_\sigma(S) u_t - H_\sigma(S) y_t] \]
\[ = E(S)^{-1} S^j G_\sigma(S)(\xi_t^j - \theta_\sigma, \xi_t) \]
\[ + E(S)^{-1} R(S)^{-1} S^{j+1} B_\sigma(S)[G_\sigma(S) u_t - H_\sigma(S) y_t] \]
\[ j = 0, 1, \ldots, p + 1, \]

and, furthermore, we have

\[ \sum_{j=0}^{p+1} (S^j \xi_t^j)^2 \leq 2 \|G_\sigma(S)^2 \sum_{j=0}^{p+1} \sum_{k=0}^{q-1} (S^j E^{-1} S^{-1})(\xi_t^j - \theta_\sigma, \xi_t)) \]
\[ + 2 \sum_{j=0}^{p+1} (E^{-1}(S)R^{-1}(S)S^{j+1} B_\sigma(S)[G_\sigma(S) u_t - H_\sigma(S) y_t])^2 \]
\[ \leq 2 \|G_\sigma(S)^2 (p + q + 1) \sum_{j=0}^{p+q} (S^j E^{-1} S^{-1})(\xi_t^j - \theta_\sigma, \xi_t)) \]
\[ (3.27) \]

and similarly, by (3.26) we get

\[ \sum_{j=0}^{q} (S^j \xi_t^j)^2 \leq 2 \|H_\sigma(S)^2 \sum_{j=0}^{q} \sum_{k=0}^{p} (S^j E^{-1} S^{-1})(\xi_t^j - \theta_\sigma, \xi_t) \]
\[ + 2 \sum_{j=0}^{q} (E^{-1}(S)R^{-1}(S)S^j A_\sigma(S)[G_\sigma(S) u_t - H_\sigma(S) y_t])^2 \]
\[ \leq 2 \|H_\sigma(S)^2 (p + q + 1) \sum_{j=0}^{p+q} (S^j E^{-1} S^{-1})(\xi_t^j - \theta_\sigma, \xi_t) \]
\[ (3.28) \]

By Lemma 3.1 and (3.24) we see that

\[ \limsup_{t \to \infty} \frac{1}{t+1} \sum_{j=0}^{p+1} \int_0^t (E^{-1}(S)R^{-1}(S)S^{j+1} B_\sigma(S)[G_\sigma(S) u_t - H_\sigma(S) y_t])^2 ds < \infty \quad \text{a.s.} \]
and that
\[
\limsup_{t \to \infty} \frac{1}{t + 1} \sum_{j=0}^{q} \int_0^t \left( E^{-1}(S)R^{-1}(S)S^j A_\sigma(S)[G_\sigma(S)u_S - H_\sigma(S)y_S] \right)^2 ds < \infty \quad \text{a.s.}
\]

Therefore, by (3.27), (3.28), and Lemma 3.1 we conclude that for some \( \nu_1 < \infty \) which is independent of \( t \)
\[
(3.29)
\]
\[
\frac{1}{t + 1} \int_0^t \left[ \sum_{k=0}^{p+1} (S^k \xi^y_s)^2 + \sum_{k=0}^{q} (S^k \xi^u_s)^2 \right] ds
\]
\[
\leq 2(p + q + 1)[\|G_\sigma(S)\|^2 + \|H_\sigma(S)\|^2] \frac{1}{t + 1} \sum_{j=0}^{p+q} \int_0^t \left[ \frac{S^j}{E(S)} (\xi^y_s - \theta_{\sigma_1}^y \zeta_s) \right]^2 ds + \nu_1
\]
\[
\leq \frac{2(p + q + 1)s_1}{2C_e(p + q + 1)} \cdot \frac{1}{t + 1} \int_0^t (\xi^y_s - \theta_{\sigma_1}^y \zeta_s)^2 ds + \nu_1
\]
\[
\leq \frac{1}{t + 1} \frac{1}{t + 1} f(\sigma_i, t) + \nu_1 \quad \text{a.s.,} \quad t \geq T_{\sigma_i},
\]
where (3.17), \( \sigma_i < \infty \), and \( \tau_i = \infty \) a.s. have been used for the last inequality.

Set
\[
\nu_2 = \nu_1 + \sup_{0 \leq \lambda \leq T_{\sigma_i}} \left\{ \frac{1}{\lambda + 1} \int_0^\lambda \left[ \sum_{k=0}^{p+1} (S^k \xi^y_s)^2 + \sum_{k=0}^{q} (S^k \xi^u_s)^2 \right] ds \right\}
\]
Then from (3.29) and (3.18) it follows that
\[
\sup_{0 \leq \lambda \leq t} \left\{ \frac{1}{\lambda + 1} \int_0^\lambda \left[ \sum_{k=0}^{p+1} (S^k \xi^y_s)^2 + \sum_{k=0}^{q} (S^k \xi^u_s)^2 \right] ds \right\} \leq \sigma_i^{-1} \frac{1}{t + 1} f(\sigma_i, t) + \nu_2
\]
\[
\leq \sigma_i^2 + \nu_2 + \sigma_i^{-1} \sup_{0 \leq \lambda \leq t} \left\{ \frac{1}{\lambda + 1} \int_0^\lambda \left[ \sum_{k=0}^{p+1} (S^k \xi^y_s)^2 + \sum_{k=0}^{q} (S^k \xi^u_s)^2 \right] ds \right\} \quad \text{a.s.,}
\]
i.e.,
\[
\limsup_{t \to \infty} \sup_{0 \leq \lambda \leq t} \left\{ \frac{1}{\lambda + 1} \int_0^\lambda \left[ \sum_{k=0}^{p+1} (S^k \xi^y_s)^2 + \sum_{k=0}^{q} (S^k \xi^u_s)^2 \right] ds \right\}
\]
\[
\leq (1 - \sigma_i^{-1})^{-1} [\nu_2 + \sigma_i^2] < \infty \quad \text{a.s.}
\]
From this and (3.14), assertion (b) follows. \( \Box \)

**Appendix A. Proof of Lemma 2.2.** Let
\[
(A.1) \quad N = \begin{bmatrix} \beta_1 & \beta_2 & \ldots & \beta_{n+1} \\ 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 1 & 0 \end{bmatrix}, \quad H = \begin{bmatrix} 1, 0, \ldots, 0 \end{bmatrix}^\top,
\]
and, for any \( t \in [T^k, T^{k+1}) \),
\[
U_t(k) = [u_t, S_k u_t, \ldots, S_k^n u_t]^\top.
\]
Then from the definition of $u_t$ it follows that for any $t \in (T^k, T^{k+1})$,
\[ \frac{dU_t(k)}{dt} = NU_t(k) \quad \text{with } U_{T^k}(k) = HL_{T^k}, \]
i.e.,
\[ S_k U_t(k) = N^{-1} U_t(k) - N^{-1} HL_{T^k} \quad \forall t \in [T^k, T^{k+1}). \]

We first to show that there exist constants $c > 0$, $\gamma > 1$ and $k_0$ such that
\[ \lambda_{\min} \left( \int_{T^k}^{T^{k+1}} (S_k U_s(k))(S_k U_s(k))^\top ds \right) \geq c \gamma^{T^{k+1}} L_{T^k}^2 \quad \forall k \geq k_0, \]
where here and hereafter $\lambda_{\min}(X)$ denotes the minimal eigenvalue of matrix $X$.

From (2.5) and (A.1) we see that the characteristic polynomial $\det(xI - N) = (x - \alpha)^{n+1}$ of $N$ coincides with the minimal polynomial of $N$. Thus there is a nonsingular $(n + 1) \times (n + 1)$ matrix $P$ such that
\[ \Lambda \triangleq P^{-1} N P = \begin{bmatrix} \alpha & 1 & \cdots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & \alpha \end{bmatrix}_{(n+1)\times(n+1)}. \]

Let $\bar{U}_t(k) = P^{-1} U_t(k)$ and $\bar{H} = P^{-1} H$. Then (A.2) is equivalent to
\[ \frac{d\bar{U}_t(k)}{dt} = \Lambda \cdot \bar{U}_t(k) \quad \text{with } \bar{U}_{T^k}(k) = \bar{H}L_{T^k} \quad \forall t \in (T^k, T^{k+1}). \]

Noting that for a given positive semidefinite matrix $U$,
\[ \lambda_{\min}(PUP^\top) \geq \lambda_{\min}(PP^\top) \lambda_{\min}(U) \]
and $\lambda_{\max}(PP^\top) \leq \|P\|^2$, by (A.3) and inequality $a^2 \geq \frac{1}{2} b^2 - (b - a)^2$ we have
\[ \lambda_{\min} \left( \int_{T^k}^{T^{k+1}} (S_k U_s(k))(S_k U_s(k))^\top ds \right) \geq \frac{1}{2} \lambda_{\min}(N^{-1}N^{-\top}) \lambda_{\min} \left( \int_{T^k}^{T^{k+1}} U_s(k)U_s^\top(k) ds \right) - \|N^{-1}\|_2 T^{k+1} L_{T^k}^2 \]
\[ \leq \frac{1}{2} \lambda_{\min}(N^{-1}N^{-\top}) \lambda_{\min}(PP^\top) \lambda_{\min} \left( \int_{T^k}^{T^{k+1}} \bar{U}_s(k)\bar{U}_s^\top(k) ds \right) - \|N^{-1}\|_2 T^{k+1} L_{T^k}^2. \]

Thus, in order to show (A.4), it suffices to prove that there exist constants $c > 0$, $\gamma > 1$, and $k_0$ such that
\[ \lambda_{\min} \left( \int_{T^k}^{T^{k+1}} \bar{U}_s(k)\bar{U}_s^\top(k) ds \right) \geq c \gamma^{T^{k+1}} L_{T^k}^2 \quad \forall k \geq k_0. \]

From (A.6) it follows that
\[ \bar{U}_t(k) = e^{\Lambda t} \bar{H}L_{T^k} \quad \forall t \in [T^k, T^{k+1}), \]
which implies that
\[
\int_{T_k}^{T_{k+1}} \tilde{U}_s(k)\tilde{U}_s^T(k)ds = L_T^2 \int_{T_k}^{T_{k+1}} e^{\tilde{A} s} \tilde{H} \cdot \tilde{H}^T e^{\tilde{A}^T s}ds
\geq L_T^2 \int_{T_{k+1}^{-1}}^{T_{k+1}} e^{\tilde{A} s} \tilde{H} \cdot \tilde{H}^T e^{\tilde{A}^T s}ds
\]
\[
= L_T^2 e^{\tilde{A}(T_{k+1}^{-1})} \left( \int_0^1 e^{\tilde{A} \tilde{s}} \tilde{H} \cdot \tilde{H}^T e^{\tilde{A}^T \tilde{s}}ds \right) e^{\tilde{A}(T_{k+1}^{-1})}
\geq L_T^2 \lambda_{\min} \left( \int_0^1 e^{\tilde{A} \tilde{s}} \tilde{H} \cdot \tilde{H}^T e^{\tilde{A}^T \tilde{s}}ds \right) e^{\tilde{A}(T_{k+1}^{-1})} e^{\tilde{A}(T_{k+1}^{-1})}.
\]  
(A.8)

Note that \((N, H)\) is controllable and hence \((\tilde{A}, \tilde{H})\) is controllable. Therefore,

\[
\lambda_{\min} \left( \int_0^1 e^{\tilde{A} \tilde{s}} \tilde{H} \cdot \tilde{H}^T e^{\tilde{A}^T \tilde{s}}ds \right) > 0.
\]  
(A.9)

Set

\[
\Sigma_t = \begin{bmatrix}
1 \\
(t-1) \\
\vdots \\
1 \\
\frac{(t-1)^n}{n!} \\
\vdots \\
(t-1) \\
1
\end{bmatrix}.
\]

Then from (A.5) we see that

\[
e^{\tilde{A}(t-1)} = e^{\alpha(t-1)} \Sigma_t.
\]  
(A.10)

It is evident that

\[
\det(\Sigma_t \Sigma_t^T) = 1 \quad \text{and} \quad \lambda_{\max}(\Sigma_t \Sigma_t^T) \leq (n+1) \sum_{i=0}^{n} (t-1)^{2i},
\]

where \(\lambda_{\max}(X)\) denotes the maximum eigenvalues of \(X\).

Thus, from the fact that \(\det(X) = \prod_{i=1}^{n+1} \lambda_i(X)\) for any \((n+1) \times (n+1)\) matrix with eigenvalues \(\lambda_i(X)\) \((i = 1, \ldots, n+1)\) it follows that

\[
\lambda_{\min}(\Sigma_t \Sigma_t^T) \geq \left[ \lambda_{\max}(\Sigma_t \Sigma_t^T) \right]^{-n} \geq \left[ (n+1) \sum_{i=0}^{n} (t-1)^{2i} \right]^{-n}
\geq (n+1)^{-2n}(t-1)^{-2n^2} \quad \forall t \geq 2.
\]

From this and (A.10) we get that

\[
\lambda_{\min} \left( e^{\tilde{A}(t-1)} e^{\tilde{A}^T(t-1)} \right) \geq e^{2\alpha(t-1)} \cdot (n+1)^{-2n}(t-1)^{-2n^2} \quad \forall t \geq 2,
\]

which together with \(\alpha > 0\), (A.9), and (A.8) implies the desired result (A.7). Therefore, (A.4) is true.
We are now in a position to prove (2.13).

Write

$$\text{Adj}(I - AS) = I + A_1 S + \cdots + A_{n-1} S^{n-1},$$

$$A(S) \triangleq \det(I - AS) = a_0 + a_1 S + \cdots + a_n S^n$$

and set

$$M = \begin{bmatrix} 0 & B & A_1 B & \ldots & A_{n-1} B \\ a_0 & a_1 & a_2 & \ldots & a_n \end{bmatrix},$$

where $I$ denotes $n \times n$ identity matrix.

Clearly, we have

$$A(S)\psi_\lambda = \begin{bmatrix} \text{Adj}(I - AS)S^2 Bu_\lambda + \text{Adj}(I - AS)(x_0 \lambda + CSu_\lambda) \\ A(S)Su_\lambda \end{bmatrix}$$

$$= MSU_t + \begin{bmatrix} \text{Adj}(I - AS)(x_0 \lambda + CSu_\lambda) \\ 0 \end{bmatrix},$$

where

$$U_t = [u_t, S_{u_t}, \ldots, S^n_{u_t}]^\tau.$$ 

Therefore, we get

$$\lambda_{\text{min}} \left( \int_{T^k}^t [A(S)\psi_\lambda][A(S)\psi_\lambda]^\tau \, ds \right)$$

$$\geq \frac{1}{2} \lambda_{\text{min}} \left( M \int_{T^k}^t (SU_s)(SU_s)^\tau \, dsM^\tau \right) - c \sum_{i=0}^n \int_0^t [s^{2i} + \|S^i w_s\|^2] \, ds.$$ 

(A.11)

Using the argument in, e.g., Zhang and Chen [1991] we can obtain

$$\int_0^t \|S^i w_s\|^2 \, ds \leq ct^{2i + 3}, \quad i = 0, 1, 2, \ldots, n,$$

and

$$\lambda_{\text{min}} \left( \int_{T^k}^t [A(S)\psi_\lambda][A(S)\psi_\lambda]^\tau \, ds \right) = \min \left\{ \lambda_{\text{min}} \left( \sum_{i=0}^n \alpha_i S^i x_\lambda^\tau \psi_\lambda \right) \right\}$$

$$\leq c \sum_{i=0}^n t^{2i+1} \lambda_{\text{min}} \left( \int_{T^k}^t \psi_\lambda^\tau \psi_\lambda \, ds \right) + c \left( \sum_{i=0}^n t^{2i+1} \right) \int_0^t \|\psi_\lambda\|^2 \, ds.$$ 

(A.13)

From (A.11)–(A.13) we have

$$\lambda_{\text{min}} \left( \int_{T^k}^t \psi_\lambda^\tau \psi_\lambda \, ds \right) \geq \frac{1}{c} \lambda_{\text{min}}(MM^\tau) \left( \sum_{i=0}^n t^{2i+1} \right)^{-1} \lambda_{\text{min}} \left( \int_{T^k}^t (SU_s)(SU_s)^\tau \, ds \right)$$

$$- c \left( \sum_{i=0}^n t^{2i+1} \right)^{-1} \left( \sum_{i=0}^{n+1} t^{2i+1} \right) - \int_0^T \|\psi_\lambda\|^2 \, ds.$$ 

(A.14)
By induction we can show that
\[ |S^i u_t - S^i_k u_t| \leq t^{i-1} \int_0^{T^k} |u_s| ds \quad \forall i = 1, 2, \ldots, \forall t \in [T^k, T^{k+1}). \]

From this we see that for any \( x \in \mathbb{R}^{n+1} \) with \( ||x|| = 1 \) and \( t \in [T^k, T^{k+1}) \),
\[ \int_{T^k}^t (x^T S U_s)^2 ds \geq \frac{1}{2} \int_{T^k}^t [x^T S_k U_s(k)]^2 ds - \int_{T^k}^t \|S_k U_s(k) - S U_s\|^2 ds \]
\[ \geq \frac{1}{2} \int_{T^k}^t [x^T S_k U_s(k)]^2 ds - \int_{T^k}^t \sum_{i=1}^{n+1} |S^i u_s - S^i_k u_s|^2 ds \]
\[ \geq \frac{1}{2} \int_{T^k}^t [x^T S U_s(k)]^2 ds - \sum_{i=1}^{n+1} \left( \int_0^{T^k} |u_s| ds \right)^2 \int_{T^k}^t s^{2(i-1)} ds \]
\[ \geq \frac{1}{2} \int_{T^k}^t [x^T S U_s(k)]^2 ds - \sum_{i=1}^{n+1} \left( \int_0^{T^k} u_s^2 ds \right) T^{2i(k+1)} \]
\[ \geq \frac{1}{2} \int_{T^k}^t [x^T S U_s(k)]^2 ds - (n+1)T^{2(n+1)}(k+1) \int_0^{T^k} u_s^2 ds, \]

which implies that
\[ \lambda_{\min} \left( \int_{T^k}^{T^{k+1}} (S U_s)(S U_s)^T ds \right) \geq \frac{1}{2} \lambda_{\min} \left( \int_{T^k}^{T^{k+1}} (S_k U_s(k))(S_k U_s(k))^T ds \right) - (n+1)T^{2(n+1)}(k+1) L_{T^k}. \]

(A.15)

This together with (A.4) and (A.14) leads to
\[ \lambda_{\min} \left( \int_{T^k}^{T^{k+1}} \psi_s \psi_s^T ds \right) \geq c^{-1} T^{-(2n+1)(k+1)} \gamma T^{k+1} L_{T^k}^2 - cT^{(2n+3)(k+1)} L_{T^k}. \]

(A.16)

With \((S + 1) z_t = \psi_t\) in mind, we get
\[ \lambda_{\min} \left( \int_{T^k}^{T^{k+1}} \psi_s \psi_s^T ds \right) = \lambda_{\min} \left( \int_{T^k}^{T^{k+1}} [(S + 1) z_s][(S + 1) z_s]^T ds \right) \]
\[ = \min \int_{T^k}^{T^{k+1}} |x^T z_s + S x^T z_s|^2 ds \]
\[ \leq 4 T^{2(k+1)} \lambda_{\min} \left( \int_{T^k}^{T^{k+1}} z_s z_s^T ds \right) + 2T^{2(k+1)} \int_0^{T^k} \|z_s\|^2 ds, \]

i.e.,
\[ \lambda_{\min} \left( \int_{T^k}^{T^{k+1}} z_s z_s^T ds \right) \geq 4^{-1} T^{-2(k+1)} \lambda_{\min} \left( \int_{T^k}^{T^{k+1}} \psi_s \psi_s^T ds \right) \]
\[ - 2^{-1} \int_0^{T^k} \|z_s\|^2 ds. \]
From this, (A.16), and the definition of $L_{Tk}$ it follows that

$$\lambda_{\min}\left(\int_{Tk}^{Tk+1} z_s z_s^T ds\right) \geq c^{-1} T^{-(2n+3)(k+1)} \gamma^{Tk+1} L_{Tk}^2 - c T^{(2n+1)(k+1)} L_{Tk},$$

which implies the desired result (2.13). □

**Appendix B.** The following lemma is based on Chen [1992].

**Lemma B.1.** If $T > 1$ is a constant, $(F, C)$ is controllable, $F \in \mathcal{F}_{\sigma}$ with $\sigma < \infty$ a.s. being a stopping time, and if

(B.1) \[
\limsup_{k \to \infty} \frac{1}{T^k} \int_0^{Tk} \|x_s\|^2 ds < \infty \quad \text{a.s.}
\]

for the system

(B.2) \[
dx_t = Fx_t dt + Cdw_t, \quad t \geq \sigma,
\]

then $F$ must be stable a.s.

**Proof.** Assume that $F^\tau$ has an eigenvalue $\lambda$ with $\Re(\lambda) \geq 0$, where $\Re(x)$ denotes the real part of a complex number $x$. Let $y$ be the corresponding eigenvector, i.e., $F^\tau y = \lambda y$. Then by (B.2) we get

$$d[\Re(y^\tau x_t) + i\Im(y^\tau x_t)] = [\Re(\lambda) + i\Im(\lambda)] [\Re(y^\tau x_t) + i\Im(y^\tau x_t)] dt + y^\tau C dw_t \quad \forall t \geq \sigma,$$

i.e.

(B.3) \[
dz_t = \begin{bmatrix} \Re(\lambda) & -\Im(\lambda) \\ \Im(\lambda) & \Re(\lambda) \end{bmatrix} z_t dt + \begin{bmatrix} \Re(y^\tau C) \\ \Im(y^\tau C) \end{bmatrix} dw_t,
\]

where $\Im(x)$ denotes the imaginary part of a complex number $x$ and

$$z_t = \begin{bmatrix} \Re(y^\tau x_t) \\ \Im(y^\tau x_t) \end{bmatrix}.$$

Using Ito’s formula, by (B.3) we obtain

$$dz_t^\tau z_t = 2z_t^\tau \begin{bmatrix} \Re(\lambda) & -\Im(\lambda) \\ \Im(\lambda) & \Re(\lambda) \end{bmatrix} z_t dt + 2z_t^\tau \begin{bmatrix} \Re(y^\tau C) \\ \Im(y^\tau C) \end{bmatrix} dw_t + \|y^\tau C\|^2 dt$$

$$= 2\Re(\lambda) \|z_t\|^2 dt + 2z_t^\tau \begin{bmatrix} \Re(y^\tau C) \\ \Im(y^\tau C) \end{bmatrix} dw_t + \|y^\tau C\|^2 dt,$$

which implies (Christopeit [1986]) that for any $\eta \in (0, 1/2)$,

(B.4) \[
\|z_t\|^2 = \|z_\sigma\|^2 + 2\Re(\lambda) \int_\sigma^t \|z_s\|^2 ds + O\left(\left(\int_\sigma^t \|z_s\|^2 ds\right)^{1/2+\eta}\right) + \|y^\tau C\|^2 (t-\sigma).
\]

Noting that (B.1) implies that

(B.5) \[
\limsup_{k \to \infty} \frac{1}{T^k} \int_\sigma^{Tk} \|z_s\|^2 ds < \infty,
\]
by (B.4) we have

(B.6) \( ||z_t||^2 = ||z_\sigma||^2 + 2\Re(\lambda) \int_\sigma^t ||z_s||^2 ds + O(t^{1+\eta}) + ||y^\tau C||^2 (t - \sigma) \quad \forall \eta \in (0, 1/2). \)

From controllability of \((F, C)\) it is easy to see that \( ||y^\tau C|| \neq 0 \), and hence from (B.6) and \( \Re(\lambda) \geq 0 \) it follows that for some \( t_0 \geq \sigma \) and \( c > 0 \),

\[ ||z_t||^2 \geq ct \quad \forall t \geq t_0, \ t_0 \text{ random}, \]

which contradicts (B.5).

Appendix C. Proof of Lemma 3.3. As in Appendix A, we can show that there exist constants \( c > 0, \gamma > 1, \) and \( k_0 \) such that

(C.1) \[ \lambda_{\min} \left( \int_{T^k} (S_k U_s(k))(S_k U_s(k))^\tau ds \right) \geq c \gamma^{T^k} L^2_{T^k} \quad \forall k \geq k_0, \]

where here and hereafter

\[ U_t(k) = [u_t, S_k u_t, \ldots, S_k^{p+q-1} u_t]^\tau. \]

Set \( W_t = y_0 t + SC(S)w_t + S^2 \eta_t \) and \( M = [M_1, M_2]^\tau \) with

\[ M_1^p \triangleq \begin{pmatrix} 0 & b_1 & \ldots & \ldots & \ldots & \ldots & b_q & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots \\ 0 & \ldots & 0 & b_1 & \ldots & \ldots & \ldots & b_q \end{pmatrix} \]

and

\[ M_2^p \triangleq \begin{pmatrix} 1 & a_1 & \ldots & \ldots & \ldots & \ldots & a_q & 0 & \ldots & 0 \\ 0 & 1 & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots \\ 0 & \ldots & 0 & 1 & \ldots & \ldots & \ldots & a_q \end{pmatrix} \]

Then from (3.1) and (3.11) it follows that

(C.2) \[ A(S) \varphi_t = MSU_t + [W_t, SW_t, \ldots, S^{p-1} W_t, \begin{pmatrix} 0, \ldots, 0 \end{pmatrix}^\tau, \]

where here and hereafter,

(C.3) \[ U_t = [u_t, S u_t, \ldots, S^{p+q-1} u_t]^\tau. \]

Notice that by Assumption A.2, for \( i = 1, 2, \ldots, p, \)

\[ \int_0^t (S^i \eta_s)^2 ds \leq \int_0^t \frac{t^{2i} - s^{2i}}{2i(2i - 1)(i - 1)!} \eta_s^2 ds \leq ct^{2i+1}. \]
Then similar to (A.11)-(A.14), by (C.2) we obtain

\[
\lambda_{\min}\left(\int_{T_k}^{t_l} \varphi_s^T \varphi_s^* ds\right) \geq \frac{1}{c} \left(\sum_{i=0}^{p} t^{2i+1}\right)^{-1} \lambda_{\min}\left(\int_{T_k}^{t_l} (SU_s)(SU_s)^T ds\right) - c \left(\sum_{i=0}^{p} t^{2i+1}\right)^{-1} \sum_{i=0}^{p+l} (t+1)^{2i+1} - \int_{0}^{T_k} \|\varphi_s\|^2 ds.
\]

(C.4)

By (C.3) and along the argument of (A.15) we get

\[
\lambda_{\min}\left(\int_{T_k}^{T^{k+1}} (SU_s)(SU_s)^T ds\right) \geq \frac{1}{2} \lambda_{\min}\left(\int_{T_k}^{T^{k+1}} (S_kU_s(k))(S_kU_s(k))^T ds\right) - (p + q)T^2(p+q)(k+1)L_T^k.
\]

This together with (C.1) and (C.4) leads to

\[
\lambda_{\min}\left(\int_{T_k}^{T^{k+1}} \varphi_s^T \varphi_s^* ds\right) \geq c^{-1}T^{-(2p+1)(k+1)}T^{k+1}L_T^2 - cT^2(p+q+l+2)(k+1)L_T^k.
\]

(C.5)

With \(F(S)\varphi_s^f = \varphi_t\) in mind, as in (A.13) we get

\[
\lambda_{\min}\left(\int_{T_k}^{T^{k+1}} \varphi_s^T \varphi_s^* ds\right) = \lambda_{\min}\left(\int_{T_k}^{T^{k+1}} (F(S)\varphi_s^f)(F(S)\varphi_s^f)^T ds\right) = \min_{\|x\|=1} \int_{T_k}^{T^{k+1}} \left|\sum_{i=0}^{l+1} f_i S_i x^T \varphi_s^f\right|^2 ds
\]

\[
\leq cT^{(2l+3)(k+1)} \lambda_{\min}\left(\int_{T_k}^{T^{k+1}} \varphi_s^T (\varphi_s^f)^T ds\right) + cT^{(2l+3)(k+1)} \int_{0}^{T_k} \|\varphi_s^f\|^2 ds,
\]

i.e.,

\[
\lambda_{\min}\left(\int_{T_k}^{T^{k+1}} \varphi_s^T (\varphi_s^f)^T ds\right) \geq cT^{-(2l+3)(k+1)} \lambda_{\min}\left(\int_{T_k}^{T^{k+1}} \varphi_s^T \varphi_s^* ds\right) - \int_{0}^{T_k} \|\varphi_s^f\|^2 ds.
\]

From this, (C.5) and the definition of \(L_T^k\) it follows that

\[
\lambda_{\min}\left(\int_{T_k}^{T^{k+1}} \varphi_s^T (\varphi_s^f)^T ds\right) \geq c^{-1}T^{-(2p+1)(k+1)}T^{k+1}L_T^2 - cT^2(p+q+l+2)(k+1)L_T^k,
\]

which implies the desired result, Lemma 3.3. □

Appendix D.

**Lemma D.1.** If \(A(S)\) and \(SB(S)\) are coprime, \(b_q \neq 0\) and

\[
\theta_{T^k} \xrightarrow[k \to \infty]{} \theta \text{ a.s.,}
\]
then there is an integer-valued $K$, possibly depending on a sample path such that for all $k \geq K$, $A_k(S)G_k(S) - SB_k(S)H_k(S) = E(S)$ is solvable with respect to $G_k(S)$ and $H_k(S)$ subject to $\partial(G_k(S)) \leq q - 1$ and $\partial(H_k(S)) = p$, and such that $\|G_k(S)\|^2 + \|H_k(S)\|^2 \leq k/(2c(p + q + 1))$.

Proof. Let

\[
\begin{align*}
(M.1) & \quad M_3^k = \begin{pmatrix}
1 & a_1 & \cdots & \cdots & \cdots & a_p & 0 & \cdots & 0 & 0 \\
0 & 1 & & & & \cdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & 1 & a_1 & \cdots & \cdots & a_p & 0 & 0
\end{pmatrix}, \\
(M.2) & \quad M_4^k = \begin{pmatrix}
0 & -b_1 & \cdots & \cdots & \cdots & -b_q & 0 & \cdots & 0 \\
0 & 0 & & & & \cdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & 0 & -b_1 & \cdots & \cdots & -b_q & 0
\end{pmatrix}, \\
(M.3) & \quad M' = [M_3, M_4], \\
He = [1, e_1, \ldots, e_{p+q}]^T.
\end{align*}
\]

Replacing $a_i$ and $b_j$ by their estimates $a_{i,T^k}$ and $b_{j,T^k}$, respectively, in (M.1)-(M.3) for $i = 1, \ldots, p$ and $j = 1, \ldots, q$, we correspondingly denote $M_3, M_4$, and $M'$ by $M_3,k, M_4,k$, and $M'_k$. Furthermore, if $M'_k$ is nonsingular, we set $\Psi_k = (M'_k)^{-1}He$.

Since $A(S)$ and $SB(S)$ are coprime and $b_q \neq 0$, we see that $M'$ given by (M.1)-(M.3) is nonsingular. Let $\Psi = (M')^{-1}He$ and $G(S) = \sum_{i=0}^{q-1} g_i S^i, H(S) = \sum_{i=0}^{p} h_i S^i$ with

\[
[g_0, g_1, \ldots, g_{q-1}, \ h_0, h_1, \ldots, h_{p}] \triangleq \Psi.
\]

Then recalling that

\[
\theta_{T^k} \to_{k \to \infty} \theta,
\]

we see that there is an integer $K' \geq 0$ such that for all $k \geq K'$, $M'_k$ is nonsingular,

\[
M'_k \to_{k \to \infty} M', \text{ and } \Psi_k \to_{k \to \infty} \Psi.
\]

Furthermore for all $k \geq K'$, if we set $G_k(S) = \sum_{i=0}^{q-1} g_i,k S^i, H_k(S) = \sum_{i=0}^{p} h_i,k S^i$ with

\[
[g_{0,k}, g_{1,k}, \ldots, g_{q-1,k}, \ h_{0,k}, h_{1,k}, \ldots, h_{p,k}] \triangleq \Psi_k,
\]

then we have $\partial(G_k(S)) \leq q - 1, \partial(H_k(S)) = p$ and

\[
A_k(S)G_k(S) - SB_k(S)H_k(S) = E(S).
\]

Noting that

\[
\Psi_k \to_{k \to \infty} \Psi
\]

we see that there exists an integer $K' \geq K'$ such that for all $k \geq K$, $\|G_k(S)\|^2 + \|H_k(S)\|^2 \leq k/(2c(p + q + 1))$. \(\square\)
REFERENCES


H. F. Chen (1990), *Continuous-time stochastic adaptive control stabilizing the system and minimizing the quadratic loss function*, Tech. report, Institute of Systems Science, Academia Sinica, Beijing.


