Products of Random Matrices in Control and Signal Processing

Jelel EZZINE.

ACS, ENIT, Campus Universitaire, BP 37 Le Belvédère, Tunis, TUNISIA
e-mail: jelel.ezzine@ieee.org

Abstract

The paper gives a review of the products of random matrices and their applications in different areas of research. It dwells mainly on two applications in control and signal processing. The one in control deals with the almost sure stabilization of jump parameter systems via the assignment of the leading Lyapunov exponent. The second one deals with a quite interesting application in adaptive filtering. The particularity of this problem is the finiteness of the alphabet dealt with in the filtering problem. Though this could be seen as a special case, it happens to be of great interest in digital communication. Indeed such a model is far more appropriate than classical ones. It is shown that due to the finiteness of the alphabet, that rigorous stability results can be achieved without any unrealistic independence assumptions.

Keywords: Products of random matrices, Almost-sure stability/stabilization, Lyapunov exponents, Linear matrix inequalities (LMI's), Adaptive filtering.

Notation

$(\cdot)^{\dagger}$	pseudo inverse of $(\cdot)$
$\|\cdot\|$	Euclidean norm of a vector, or induced Euclidean norm of a matrix
$\mathcal{A}$	set of matrices A_i, $i = 1, \cdots, N$
$A_i^{[l]}$	product of l matrices $A_i \in \mathcal{A}$, referred to as l-block
$\mathcal{A}^{[l]}$	set of matrices $A_i^{[l]}$, $i = 1, \cdots, N^l$
$\mathbf{E}$	expected value operator
λ_i	ith Lyapunov exponent
λ_1	leading or largest Lyapunov exponent
$\mathcal{N}$	$\{i \mid i = 1, \cdots, N\}$
$\mathcal{N}^{[l]}$	$\{i \mid i = 1, \cdots, N^l\}$
p_i	steady state probability of the matrix A_i
$p_i^{[l]}$	steady state probability of the l-block matrix $A_i^{[l]}$

Acronym	**Meaning**
a.s.	almost-surely (or with probability one)
DTR	discrete-time random dynamical system
FSMC	finite states Markov chain
FKT	Furstenberg-Kesten theorem
GEVP	Generalized Eigenvalues Problem
LMI	Linear Matrix Inequality
NSGASS	non-switching gains a.s. stabilizable
SGASS	switching gains a.s. stabilizable
SGLASS	switching gains l-block a.s. stabilizable
SGSASS	switching gains strongly a.s. stabilizable
SGSS	switching gains strongly stabilizable

1 Introduction

The product of random matrices has attracted the attention of mathematicians [1], physicists [2]and engineers in recent years[3][4][5][6][7][8]. One of the early results was proposed by Furstenberg and Kesten in 1960 [9]. In their contribution they gave the first quite general result on the almost sure stability of the products of random matrices. Their results is based on the Lyapunov exponent concept [10]. The most general result to date, which made the one century old ideas of Lyapunov applicable, is the work of Osceledest that appeared in 1968 [11]. It goes without saying that the products of random matrices found applications beyond those cited above. For instance the products of random matrices appears in disordered dynamical system[12], Ising models with quenched disorder, wave propagation through random medium, particle floating on turbulent fluid surface, discrete-time Schrödinger equations describing a spectrum of physical phenomena evolving in random potentials [13], demographic [1] and econometric [1] applications just to cite a few.

¿From an engineering application view, products of random matrices can be found on a large spectrum of situations. Systems of this type can be used to model networks with periodically varying switches, synchronously switched linear systems, randomly sampled-data systems, systems subject to failures, manufacturing systems, macroeconomics models, and last but not least to approximate nonlinear stochastic systems. References to these applications can be found in [3] and [5] and references therein.

2 The General Setup

In this section we introduce the general setup that all discussed problems will pertain to. Moreover, a key theorem is stated to facilitate the introduction of the results discussed in this paper as well as to make the presentation as self contained as possible. The second section discusses the new and important issues of computational complexity related to the products of random matrices.

2.1 Discrete-Time Random Dynamical Systems

In this section, attention is restricted to linear discrete-time systems with randomly varying parameters. This class of systems will be called linear discrete-time random (DTR) dynamical systems in contrast to linear discrete-time stochastic systems where randomness enters the system as an input and/or measurement noise. N-form hybrid systems are a special class of DTR dynamical systems.

More precisely, this section deals with the following problem : Let $\{A_i, i \in \mathbf{N}\}$ be a sequence of random, $n \times n$ matrices. To each $x_0 \in \mathbf{R}^n$ associate the process $\{X_k, k \in \mathbf{N}\}$ with values in $\mathbf{R}^n$, which is the solution to

$$X_{k+1} = A_k X_k, \quad k \in \mathbf{N}, \quad X_0 = x_0. \tag{1}$$

We have $X_{k+1} = A_k \cdots A_1 x_0$. The asymptotic behavior of this process is addressed by the following theorem.

Furstenberg-Kesten Theorem (FKT) *[9] : Let $\{A_i, \;\; i \in \mathbf{N}\}$ be a stationary, metrically transitive stochastic process with values in the set of $n\times n$ matrices such that $\mathbf{E}\{\log^+ \|A_0\|\} < \infty$. Then, with probability one*

$$\begin{aligned} \lim_{k\to\infty} \frac{1}{k}\mathbf{E}\{\log \|S_k\|\} &= \inf_k \frac{1}{k}\mathbf{E}\{\log \|S_k\|\} \qquad (2) \\ &= \lim_{k\to\infty} \frac{1}{k} \log \|S_k\| := \lambda_1, \end{aligned}$$

where $S_k := A_k \cdots A_1$ and $\lambda_1 \in \mathbf{R} \cup \{-\infty\}$, is the largest Lyapunov exponent of the process.

The FKT is a generalization of Birkoff's ergodic theorem to the case of matrix valued functions. It is also a simple version of the Osceledest theorem [30]. Though it provides a necessary and sufficient a.s. stability criterion, it is highly impractical from a design point of view. Therefore, one way to avoid this difficulty is to use the FKT to derive simpler criteria to test for the a.s. stability of the system, as well as aid in devising controllers.

2.2 Products of Random Matrices and Computational Complexity

Since the products of random matrices started getting more and more attention, and since their applications kept multiplying, some rigorous and definitive answers were called for as to the computational complexities related to the products of random matrices such as computing their Lyapunov exponents. It turned out, just lately, that such a task is highly complex and that it is highly unlikely that efficient algorithms will be found to tackle such problem and related ones. Indeed, the ability to compute arbitrarily accurate approximations of the Lyapunov exponent and related quantities does not rule out that the problem of deciding wether the underlying dynamical system is stable or not is undecidable. A plethora of related computational problems is reviewed in [14].

3 Leading Lyapunov Exponent Bounds

There are few results concerning the approximation of the leading Lyapunov exponent. Most of these approximations can be found mainly in the physics literature, for instance the reader can refer to [15], [16], [17], and references therein. An information theoretic based approximation of the Lyapunov spectrum can be found in [8]. There are also results dealing with the difficult and important problem of stopping rules. In [7] an asymptotically exact probabilistic stopping rule is proposed. It is important to mention that it is the first rigorous stopping rule, though expected to be quite conservative.

3.1 The Furstenberg-Kesten Bounds (FKB)

A series of increasingly tighter upper bounds on the largest Lyapunov exponent are provided along with a simple lower bound. These bounds provide a motivation for the main result of the paper which will be introduced later.

Theorem 1: *The largest Lyapunov exponent, λ_1, of a homogeneous N-form DTR dynamical system, with a stationary irreducible FSMC satisfies the following inequality*

$$\begin{array}{ll} \frac{1}{n}\sum_{i\in\mathcal{N}} p_i \log|det(A_i)| & \leq \lambda_1 \\ \qquad\quad \lambda_1 & \leq \sum_{i\in\mathcal{N}^{[l]}} p_i^{[l]} \log\|TA_i^{[l]}T^{-1}\| \\ & := \lambda^{[l]}, \end{array} \tag{3}$$

where T is any similarity transformation.

Moreover,

$$\lambda^{[ql]} \leq q\lambda^{[l]}, \quad q,l \in \mathbf{N}\setminus\{0\}. \tag{4}$$

It is important to note that the lower bound holds for the Lyapunov spectrum and not only for λ_1. That is, the lower bound is a lower bound for the smallest Lyapunov exponent as well. Consequently, the lower and upper bounds estimate the spread of the Lyapunov spectrum, thus providing valuable information regarding the slowest as well as fastest dynamics.

Though the first part of the proposition asserts that the proposed sequences of upper bounds eventually converge to the largest Lyapunov exponent, they shed no light on the *nature* of this convergence! For instance, it would be of practical interest to identify, at least, one *monotonically* convergent subsequence. This subsequence would allow one, to systematically pick increasingly tighter bounds. For example, the following is such subsequence: $\lambda^{[q^{i+1}l]} \leq \lambda^{[q^i l]}$, $i \in \mathbf{N}$. Thus, it is possible to pick as large l as one needs, to achieve a desired accuracy while computing λ_1.

The following corollaries provide useful results for stabilization purposes. The first supplies a sufficient robust a.s. stability criterion for parametrically perturbed hybrid systems. The second furnishes simple sufficient a.s. stability criteria, as well as a necessary a.s. stability test.

Corollary 1: *The N-form hybrid system with a stationary irreducible FSMC is a.s. stable if*

$$\sum_{i \in \mathcal{N}} p_i \|TA_iT^{-1}\| < 1, \tag{5}$$

where T is a similarity transformation, and is a.s. stable unless

$$\prod_{i \in \mathcal{N}} |det(A_i)|^{p_i} < 1. \tag{6}$$

3.2 Simultaneous Norm Minimization by Scaling

Obviously, for an arbitrary T, the above upper bounds are in general quite conservative. In order to alleviate this shortcoming, we use the degree of freedom provided by T, to get tighter bounds. Therefore, consider a single similarity transformation $z = Tx$, which is applied to all systems. The new systems will be

$$z(n+1) = TA_iT^{-1}z(n). \tag{7}$$

We than have

$$e^{\lambda_1} \leq \sum_{i \in \mathcal{N}} p_i \|TA_iT^{-1}\|. \tag{8}$$

The next step is to minimize the latter upper bound over nonsingular matrices T. In the sequel an LMI based solution of the latter optimization problem is proposed.

Corollary 2: *The λ_1 of a homogeneous N-form hybrid system with a stationary irreducible FSMC satisfies the following inequality*

$$e^{\lambda_1} \leq \sum_{i \in \mathcal{N}} p_i \alpha_i, \tag{9}$$

where the optimal α_i's are given by the following linear optimization problem subject to LMI constraints

$$\begin{array}{l} minimize \sum_{i \in \mathcal{N}} p_i \alpha_i \\ subject\ to \quad P = P^T > 0, \;\; A_i^T P A_i < \alpha_i^2 P, \\ i = 1, \cdots, N. \end{array} \tag{10}$$

where the p_i's are the steady state probabilities of the FSMC.

Unfortunately, the above mathematical program is not an LMI. Nevertheless, it can still be solved via a series of LMI's, where one would apply a bisection method in the n-dimensional space of the α_i's. While this approach will yield the optimal solution, it is computationally demanding. Thus, as a compromise between optimality and computational burden, a suboptimal solution, based on a series of LMI's is proposed in the next section.

This is a useful result, for it solves the robust stability problem for the linear difference inclusion (LDI) defined by the N forms of the hybrid system.

3.3 An Approximate Solution

In this section we propose an approximate solution to the optimization problem defined in corollary 4.3. The problem is broken in a series of LMI's which are solved iteratively.

Scaling Algorithm (SA):

1. $k = 0$; set $P = I$
2. solve the following LMI's for fixed P

$$\begin{array}{ll} minimize & \sum_{i \in \mathcal{N}} p_i \alpha_i \\ subject\ to & A_i^T P A_i \leq \alpha_i^2 P, \quad i = 1, \cdots, N. \end{array} \tag{11}$$

3. solve the following GEVP's using the α_i's of the previous step

$$\begin{array}{ll} minimize & t \\ subject\ to & A_i^T P A_i \leq t\alpha_i^2 P, \quad i = 1, \cdots, N. \end{array} \tag{12}$$

4. go to the second step, or stop if there are no further improvements.

4 Two Novel Applications in Control and Signal Processing

The two applications that are presented in this section, draw on two research activities pursued by the author and his colleagues. The first one deals with the almost sure stabilization of Linear Jump Parameter Systems (LJPS). The problem, since related to the products of random matrices, remains unsolved and some approximate means are introduced. The second tackles the problem of adaptive filtering in the highly practical case of finite alphabet. Though this is a special case it models digital communication systems. The breakthrough in respect to adaptive filtering is the complete elimination of the wrong independence assumption that plagues this important DSP area.

4.1 Stabilization of Jump Linear Systems

The discrete-time hybrid systems considered in this paper are assumed to have the form

$$\left\{ \begin{array}{lcl} x(k+1) & = & A(r(k))x(k) + B(r(k))u(k) \\ y(k) & = & C(r(k))x(k) \end{array} \right. \tag{13}$$

where x is the system state vector of dimension n, u is the control input vector of dimension p, y is the output vector of dimension m, and $r(k)$ is the *form index* which is either a deterministic or a stochastic scalar sequence taking values in the finite index set $\mathcal{N} = \{1, 2, ..., N\}$. The system takes the realization $\sum_i = (A_i, B_i, C_i)$ when $r(k) = i$, with $i \in \mathcal{N}$. This realization is called the ith *form*. The primary emphasis of the work is the case where $r(k)$ is a stochastic process. In this case $r(k)$ is assumed to be a Finite State Markov Chain (FSMC) with transition probabilities

$$Prob.\{r(k+1) = j | r(k) = i\} = p_{ij}. \qquad (14)$$

It is assumed, unless stated otherwise, that the FSMC is stationary and irreducible.

All available stabilization algorithms, for hybrid systems, can be mostly classified into two classes. While the first class considers perfectly observed form and state processes, the second considers unobserved or unknown ones. These stabilization algorithms achieve mean-square stability, to the exception of those reported in where a.s. stabilization is achieved directly through the negativity of the Lyapunov spectrum.

4.1.1 On Switching Gains Stabilization

The stabilization of hybrid systems was almost always carried via optimal control tools. While this route yields satisfactory results, it fails short from shedding light on the mechanism(s) taking place while achieving stabilization. Moreover, it is a mean-square stabilization approach. In this section we will formulate an a.s. stabilization problem, based on the above proposed bounds.

We are concerned with the following problem. Given N desired eigenstructures, or desired Lyapunov spectrum, specified through desired matrices A_{d_i}, $i = 1, \cdots, N$, find K_i, $i = 1, \cdots, N$, and a similarity transformation T such that

$$\|A_{d_i} - (A_i - B_i K_i)\|, \;\; i = 1, \cdots, N, \qquad (15)$$

are simultaneously minimized, and the left hand side of the following inequality

$$\sum_{i \in \mathcal{N}} p_i \|T(A_i - B_i K_i)T^{-1}\| < 1, \qquad (16)$$

is smallest.

Should the above goals be achieved, the closed-loop system is made a.s. exponentially stable via the switching gains K_i. Replacing the K_i's by K, will result in a constant or a non-switching gain controller.

4.1.2 Almost-Sure Stabilizability

Whereas a.s. stability is a more practical notion than mean-square stability, and since the latter has long appeared in the literature, the former one is defined here. The natural way to define a.s. stabilizability is through the negativity of the leading Lyapunov exponent, in complete analogy with eigenvalues negativity.

Definition 1: *A hybrid system (1, 2), with an irreducible FSMC, is said to be switching gains almost-surely stabilizable (SGASS), via static state feedback, if there exist K_i, $i = 1, \cdots, N$, such that $\lambda_1 < 0$. It is said non-switching gains almost-surely stabilizable (NSGASS) if $K_i = K, i = 1, \cdots, N$.*

Theorem 2: *A hybrid system (1, 2), with an irreducible FSMC, is SGASS, if there exist $l > 0$, a similarity transformation T, and K_i, $i = 1, \cdots, N$, such that*

$$\sum_{i \in \mathcal{N}^{[l]}} p_i^{[l]} \log \|T(A_i - B_i K_i)^{[l]} T^{-1}\| \leq 0. \qquad (17)$$

It is NSGASS if the above inequality holds for $K_i = K$, $i = 1, \cdots, N$.

Motivated by the above result, the following modified stabilization definition is introduced.

Definition 2: *A hybrid system (1, 2), with an irreducible FSMC, is said to be switching gains l-block a.s. stabilizable (SGLASS), via static state feedback, if there exist a similarity transformation T, and gains K_i, $i = 1, \cdots, N$, such that*

$$\sum_{i \in \mathcal{N}^{[l]}} p_i^{[l]} \log \|T(A_i - B_i K_i)^{[l]} T^{-1}\| \leq 0, \qquad (18)$$

and switching gains strongly a.s. stabilizable (SGSASS), or switching gains strongly stabilizable (SGSS), if it is 1-block switching gains almost-surely stabilizable.

In order to check if the system is SGSASS, and compute the N gains that will achieve a.s. stabilization, via switching gains, along with the norm scaling T matrix, in case it is SGSASS, one simply applies the following result.

Theorem 3: *A hybrid system (1, 2), with an irreducible FSMC, is (SGSS), provided that the matrices $P := T^T T > 0$, and K_i, $i = 1, \cdots, N$, generated by the following mathematical program*

$$\begin{array}{ll} \textit{minimize} & \sum_{i \in \mathcal{N}} p_i \alpha_i \\ \textit{subject to} & P = P^T > 0, \\ & (A_i - B_i K_i)^T P (A_i - B_i K_i) < \alpha_i^2 P, \\ & i = 1, \cdots, N. \end{array} \qquad (19)$$

satisfy

$$\sum_{i \in \mathcal{N}} p_i \log \|T(A_i - B_i K_i)T^{-1}\| \leq 0. \qquad (20)$$

It is NSGSS if the above inequality holds for $K_i = K$, $i = 1, \cdots, N$.

An approximate solution to the nonlinear optimization problem, via LMIs [18], can be obtained as follows.

Test Algorithm (TA):

1. $k = 0$; set $P = I$

2. solve the following LMI's for fixed P

$$\begin{array}{l} \underset{K_i,\ \alpha_i}{minimize}\ \ \alpha_i \\ subject\ to \\ \begin{bmatrix} -\alpha_i P & (A_i - B_i K_i) \\ (A_i - B_i K_i)^T & -\alpha_i P^{-1} \end{bmatrix} \le 0, \\ i = 1, \cdots, N. \end{array} \quad (21)$$

3. solve the following GEVP's using the K_i's and α_i's of the previous step using the SA Algorithm.

$$\begin{array}{l} minimize\ \ t \\ subject\ to\ \ (A_i - B_i K_i)^T P (A_i - B_i K_i) \le t\alpha_i^2 P, \\ i = 1, \cdots, N. \end{array} \quad (22)$$

4. go to the second step, or stop if there are no further improvements.

The stabilization algorithm proposed in the sequel provides either switching or non-witching gains controllers, as the difference to achieve one or the other is straight forward.

4.1.3 Stabilization Algorithms

For controller design, it is customary to have desired design specifications. For instance eigenstructure assignment is a classical design specification. The analog of eigenstructure assignment, or at least eigenvalues assignment, is the Lyapunov spectrum assignment. However, this design concept, for jump parameters systems, or any other class of systems besides LTI and periodic ones, remains a challenging open problem. One way to alleviate this lack of a design procedure, to assign the Lyapunov spectrum, is to impose "similar" specifications through other means. For instance, one can require that the closed loop matrices be as close as possible, in some sense, to N desired matrices A_{d_i}. This way the resulting gains, will try to achieve a.s. stability as well as making the resulting closed loop matrices close to the desired ones in the sense specified by the design algorithm. Should the gains succeed in exactly matching the desired matrices, the closed-loop hybrid system would inherit the desired hybrid system's dynamical characteristics such as its Lyapunov spectrum.

4.1.4 The LMI Approach to Almost-Sure Stabilization

Due to its mathematical tractability, the Frobenious norm was usually used as a measure of closeness between matrices. However, the Frobenious norm is not an operator norm, and more meaningful mean to define closeness is via the induced Euclidean norm, i.e.,

$$\|A_{d_i} - (A_i - B_i K_i)\|.$$

The optimal gain achieving the minimum of the above norm is treated in the following section as a GEVP [18]. As argued earlier, this LMI approach has the advantage of providing efficient numerical means to compute the necessary gains, as well as norm scaling transformations to improve the tightness of the sufficient stability test.

GEVP-Based Design In order to compute the N gains that will achieve a.s. stabilization, given desired closed-loop matrices A_{d_i}, along with the norm scaling matrix P, one solves separately, the two optimization problems given below.

Design Algorithm (DA):

1. solve the following using $P = I$,

$$\begin{array}{l} \underset{K_i}{minimize}\ \ \alpha_i \\ subject\ to \\ \begin{bmatrix} -\alpha_i I & (A_{d_i} - (A_i - B_i K_i)) \\ (A_{d_i} - (A_i - B_i K_i))^T & -\alpha_i I \end{bmatrix} \\ \le 0,\ i = 1, \cdots, N. \end{array} \quad (23)$$

2. to compute the similarity transformation, solve the following with the K_i's found in the first step, using SA.

$$\begin{array}{l} minimize\ \ \sum_{i \in \mathcal{N}} p_i \gamma_i \\ subject\ to\ \ (A_i - B_i K_i)^T P (A_i - B_i K_i) < \gamma_i^2 P, \\ i = 1, \cdots, N. \end{array} \quad (24)$$

Note that the outcome of the first optimization step might not yield a negative upper bound and still achieve a.s. stability. However, computing tighter bounds, as done in the second optimization step, or better the resulting Lyapunov exponent is the only way to tell if the system has been stabilized or not. However, the computation of the Lyapunov exponent is computationally much more demanding.

Numerical applications can be found in the cited references.

4.2 Convergence of Adaptive Filters

4.2.1 Background

Due to its simplicity, the LMS algorithm (or its variants: normalized LMS and sign algorithms such as clipped algorithm, pilot algorithm and zero-forcing algorithm) is widely used for adaptive signal processing. In a general context, a rigorous convergence analysis is difficult and has been hard to find.

The LMS algorithm consists in calculating recursively the parameter vector H^{opt} that optimizes, in mean square sense, the linear estimation of a measured random signal a_k, from an observation vector X_k having the same dimension d as H. It is given by:

$$H_{k+1} = H_k + \mu X_k (a_k - X_k^T H_k) \quad (25)$$

where the step size μ is a positive constant and X^T denotes the transpose of the vector X. Equivalently (1) becomes:

$$V_{k+1} = (I - \mu X_k X_k^T) V_k + \mu X_k b_k \quad (26)$$

Using the filter parameter deviation vector from the optimal ones $V_k = H_k - H^{opt}$ and the optimal filter output noise b_k.

The dynamical behavior of the algorithm is related to the transition matrix U_{kq}:

$$U_{kq} = \prod_{i=k}^{q} (I - \mu X_i X_i^T). \tag{27}$$

To overcome the complexity of this analysis due to the random nature of U_{kq}, an average approach is usually considered. It deals with the behavior study of both $E(V_k)$ and $E(V_k V_k^T)$. Since, X_k and V_k are dependent, it is difficult to establish a relation between $E(V_{k+1} V_{k+1}^T)$ and $E(V_k V_k^T)$, in fact we have :

$$\begin{aligned} E(V_{k+1} V_{k+1}^T) &= E\left((I - \mu X_k X_k^T) V_k V_k^T (I - \mu X_k X_k^T)^T\right) \\ &\quad + \mu^2 E(b_k^2) E(X_k X_k^T). \end{aligned} \tag{28}$$

where b_k is a second-order zero mean sequence, independent of X_k.

Different approaches are proposed to study the convergence of the LMS in the quadratic mean sense and almost sure sense. They can be classified in three categories:

- $C1$ Direct approaches: it consists in studying the behavior of the transition matrix U_{kq} .
- $C2$ Approximation approaches: it deals with stochastic approximation methods, such as ODE or perturbation expansion methods .
- $C3$ Simplified approaches: it deals with a particular model of the input sequence, e.g., ϕ-mixing model or one that decouples the distribution of X_k in a radial and angular distributions .

These approaches need generally the following assumptions:

- $H0$ Ergodicity of X_k,
- $H1$ Excitation condition of the algorithm: e.g., invertibility of the input covariance matrix $E(X_k X_k^T)$,
- $H2$ Boundedness of X_k or of its moments,
- $H3$ Independence: usually, the unrealistic but helpful assumption of the statistical independence of X_k and X_{k-1} is made . Sometimes, the more realistic M-independence assumption is used .

The assumptions $H0$ and $H1$ are fundamental in identification theory, and $H2$ is realistic in many applications. The independence assumption $H3$ is mainly a technicality which is required to simplify the convergence proofs. A rigorous analysis without this assumption has been difficult to find .

In this section, we focus our study on the convergence analysis in a general digital transmission context. In this context, all studies of the algorithm's convergence are carried out, using the previous convergence results ($C1, C2$ and $C3$ approaches), thus they have the same limitations. As a way to overcome these limitations, an original and rigorous approach [6], inspired from control tools[3], adapted to the finite alphabet case encountered in digital transmission contexts. Therefore, a tailored framework is made available to study applications such as speech coding, echo cancellation, channel equalization.

4.2.2 The Finite Alphabet Case

In digital transmission context, X_k remains in a finite alphabet set $A = \{W_1, W_2, ..., W_N\}$. For example in QAM modulation context with two states $\{\pm 1\}$, when the dimension of X is $d = 2$, the finite alphabet set is :

$$A \in \left\{ \begin{pmatrix} +1 \\ +1 \end{pmatrix}, \begin{pmatrix} +1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ +1 \end{pmatrix} \right\}.$$

So we suppose that $X_k = W_{\theta(k)}$, where $\{\theta(k) : k \in \mathbf{Z}^+\}$ is a discrete-time Markov chain with finite state space $\{1, 2, ...N\}$ and transition probability matrix $P = [p_{ij}]$.

To prove almost sure and quadratic mean convergences of the LMS and its variants, we consider the following assumptions:

4.2.3 Assumptions

- $A0$ $\{\theta(k) : k \in \mathbf{Z}^+\}$ is an aperiodic Markov chain.
- $A1$ The alphabet $\{W_1, W_2, ..., W_N\}$ generates the space $\Re^d$.

It is interesting to note, that trough this approach the needed assumptions are subset of the usual ones. The assumption $A0$ is similar to $H0$. It means that $\{X_k\}$ is an ergodic sequence, and $A1$ is an excitation condition similar to $H1$. The boundedness of X_k ($H2$-assumption) is a natural outcome in digital transmission contexts. Most importantly, in the finite alphabet case, no independence assumption such as the usual $H3$ is needed.

4.2.4 Almost sure and Quadratic convergences

To prove almost sure convergence of the LMS algorithm, it is possible to split V_k in two additive terms, called transient deviation V_k^t, and fluctuation deviation V_k^f. The transient deviation V_k^t is defined by:

$$\begin{aligned} V_{k+1}^t &= (I - \mu X_k X_k^T) V_k^t = \prod_{i=0}^{k} (I - \mu X_i X_i^T) V_0^t \\ &= \prod_{i=0}^{k} (I - \mu W_{\theta(i)} W_{\theta(i)}^T) V_0^t = \left(\prod_{i=0}^{k} A_{\theta(i)}\right) V_0^t. \end{aligned} \tag{29}$$

where $V_0^t = V_0$, and depends only on the initial conditions V_0, but not on the output noise b_k. When the algorithm works properly, V_k^t should forget the initial deviation value V_0 and decreases to zero almost surely.

Theorem 4: *If the assumptions $A0$ and $A1$ hold, then V_k^t converge to zero with probability 1 as well as in the quadratic mean sense.*

It is interesting to note that the proof necessitates only elementary algebra [6]. The condition on μ is quite simple and useful in digital transmission contexts. In the noiseless case ($b_k = 0$), V_k as V_k^t converges almost surely to zero. When b_k is almost surely bounded V_k is almost surely bounded too.

The quadratic mean convergence of the LMS in the finite alphabet case is carried through, without using the constraining independence assumption $H3$. As in the previous section, this analysis can be extended to the complex case as well as to clipped algorithms.

The convergence analysis presented deals with the real version of the LMS, i.e., a_k, X_k and H_k are real-valued. However, it can be extended to the complex case as well as to the normalized LMS, clipped algorithm and zero-forcing algorithm, without any additional difficulties.

5 Conclusion

The paper gives a review of the products of random matrices and their applications in different areas of research. It dwells mainly on two applications in control and signal processing. The one in control deals with the almost sure stabilization of jump parameter systems via the assignment of the leading Lyapunov exponent. These are only preliminary results for this important unsolved problem. The second part of the paper deals with a quite interesting application in adaptive filtering. The particularity of this problem is the finiteness of the alphabet dealt with in the filtering problem. Though this could be seen as a special case, it happens to be of great interest in digital communication. Indeed such a model is far more appropriate than classical ones. It turned out, due to the finiteness of the alphabet, that rigorous stability results can be achieved without any unrealistic independence assumptions. Both of the reviewed topics are in their infancy, especially the second one, and this paper hopes to ignite interests in both of these important and difficult research areas.

References

[1] J. Cohen, H. Kesten, and C. M. Newmann, (Eds.) *Random matrices and their applications*, AMS, Providence, RI, 1984.

[2] A. Crisanti, G. Paladin, and A. Vulpiani, Products of random matrices in statistical physics, Springer Series in Solid States Sciences, Springer Verlag, New York, 1991.

[3] J. Ezzine, and A. H. Haddad, "On the controllability and observability of hybrid systems," *Int. J. Contr.*, Vol. 49, No. 6, pp. 2045-2055, 1989.

[4] J. Ezzine, and A.H. Haddad, "On the stabilization of two-form hybrid systems via averaging," *Proc 22nd Annual Conf. on Information Sciences and Systems*, pp. 579-584, Princeton, 1988.

[5] J. Ezzine, and A.H. Haddad, "Error bounds in the averaging of hybrid systems," *IEEE Trans. Contr.*, Vol. 34, No. 11, pp. 1188-1192, November 1989.

[6] H. Besbes, M. Jaidane-Saidane and J. Ezzine, " Convergence des filtres adaptatifs dans le contexte alphabet fini." To be apper in GRETSI 1997, France.

[7] J. Ezzine, "On a practical stopping rule for the numerical computation of the Lyapunov spectrum," *Proc. IEEE 11th ACC,* pp. 1057-1059, Chicago, IL, June 1992.

[8] J. Ezzine, "An entropic approach to the numerical computation of the Lyapunov spectrum," *14th IMACS World Congress,* Georgia Tech., Atlanta, GA, 1994.

[9] H. Furstenberg, and R. Kesten, "Products of random matrices," *Ann. Math. Stats.*, Vol. 31, pp. 457-469, 1960.

[10] A.M. Lyapunov, "Probléme général de la stabilité du mouvement," 1982, Reprint *Ann. of Math. Studies*, Vol. 17, Princeton : Princeton Univ. Press, 1949.

[11] V.I. Oseledec, "A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems," *Trans. Moscow Math. Soc.*, Vol. 19, pp. 197-231, 1968.

[12] B. Derrida, and E. Gardner, "Lyapunov exponent of the one dimentional Anderson model: weak disorder expansions," *J. Physique*, No. 45, pp.1283-1295, 1984.

[13] B. DErrida, K. Mecheri, and J. L. Pichard, "Lyapunov exponents of products of random matrices: weak disorder expansion. Application to locolisation," *J. Physique*, No. 48, pp. 733-740, 1987.

[14] V. D. Blondel and J. N. Tsitsiklis, "A survey of computational complexity results in systemsn and control," *Automatica*, No. 36, pp. 1249-1274, 2000.

[15] J. M. Deutsch, and G. Paladin, "Product of random matrices in a microcanonical ensemble," *Physical Review Letters*, Vo. 62, No. 7, pp. 695-699, 1989.

[16] R. Mainieri, "Cycle expansion for the Lyapunov exponent of a product of random matrices," *Chaos,* Vol. 2, No. 1, pp. 91-97, 1992.

[17] R. Mainieri, "Zeta function for the Lyapunov exponent of a Product of random matrices ," *Physical Review Letters*, Vo. 68, No. 13, pp. 1965-1968, 1992.

[18] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in systems and control theory,* Philadelphia, Siam Studies 15, 1994.