COMMUNICATIONS IN INFORMATION AND SYSTEMS
Vol. 11, No. 4, pp. 327-342, 2011

002

SWITCHED LINEAR SYSTEMS: STABILITY AND THE CONVERGENCE OF RANDOM PRODUCTS

B. HANLON*, N. WANG†, M. EGERSTEDT‡, AND C. MARTIN§

Abstract. In this paper we provide conditions for the stability of discrete time switched linear systems. We accomplish this by calculating the mean and covariance of the set of matrices obtained by using all possible switching sequences. The theory of switched linear systems has received considerable attention in the systems theory literature in the last two decades. However, for discrete time switched systems the literature is much older going back to at least the early 1960's with the publication of the paper of Furstenberg and Kesten in the area of products of random matrices, or, if you like, the random products of matrices. The way that we have approached this problem is to consider the switched linear system as evolving on a partially ordered network that is, in fact, a tree. This allows us to make use of the developments of 50 years of study on random products that exist in the statistics literature.

1. Introduction. The theory of switched linear systems has received considerable attention in the systems theory literature in the last two decades and there is a wealth of solid results concerning the stability and stabilizability of such systems. A recent survey by Lin and Antsaklis, [14], gives a rather complete overview of the field. The main theorem cited there on stability for discrete time switched linear systems, (their Theorem 6, [13]), requires that 2^n matrix norms must be calculated. In this paper we give a simpler calculation but one that will only hold for "almost all" [5] switching sequences. An excellent and readable source for the general theory of switched linear systems is the monograph by Liberzon, [12]. In his book the primary focus is on continuous dynamics but the concepts are much the same. However, for discrete time switched systems the literature is much older than the systems theory literature, going back to at least the early 1960's with the publication of the papers of Furstenberg and Kesten, [9, 10], in the area of products of random matrices, or, if you like, the random products of matrices. In the statistical literature on random products "almost all" is the usual condition.

We approach the problem in the following way. We study the switched linear system as evolving on a partially ordered network that is in fact a tree. This allows us to make use of the developments of 50 years of study on random products that exist in the statistics literature.

*Department of Statistics, University of Wisconsin, Madison, WI 53706-1510, E-mail: hanlon@stati.wisc.edu

†Department of Statistics, Virginia Polytechnic Institute and State University Blacksburg, VA 24061-0002, E-mail: nwang29@vt.edu

‡School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia, 30332, E-mail: magnus.egerstedt@ece.gatech.edu

§Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX, 79409-1024, USA, E-mail: clyde.f.martin@ttu.edu

The motivation for restudying this problem came from reading a series of papers by S.A. Murphy, [16, 17, 18], on dynamic treatment regimes and associated dynamic clinical trials. In fact the thesis by Wang, [22], was directly related to this topic. The central idea behind dynamic treatments is that the treatments are switched between two or more treatments types in hopes of producing an improvement over any single treatment. A natural model for such a course of treatments lies in the area of switching systems.

We extend the paper [23] by specifically studying the stochastic behavior of the products. We characterize the convergence in term of probabilistic convergence and give a series of examples to show that we do not have convergence for all possible switching sequences; instead, we have convergence for all but a set of measure zero. We begin with the stochastic preliminaries. Our goal is to prove the sequence $\mathbf{x}_n$ converges almost surely, for example $\mathbf{x}_n \xrightarrow{a.s.} \mathbf{0}$, would say the sequence converges to $\mathbf{0}$ except for possibly on a set of measure zero. In fact, this result provides an indirect proof that the de-stabilizing sequence found in our counterexamples belongs to a set of measure zero. Using standard probabilistic arguments, we reduce the problem of almost sure convergence of the random vector to studying the mean and variance of each component.

Almost sure stability has been studied in the monograph by Costo, et al [4] and there the stability is determined in terms of the spectral radius of a certain set of matrices. Similar results are obtained by Dai et al [7] and in fact in that paper deep results are obtained in terms of Lyapunov exponents. In this paper the conditions for convergence are obtained in terms of the eigenvalues of a pair of matrices. The computation of the eigenvalues is somewhat simpler than the spectral radius calculations and the calculations for the Lyapunov exponents. The calculations for both the spectral radius and Lyapunov exponent are known to be hard if not impossible, [21]. Ogura and Martin, [19], have shown that the results on the spectral radius and the results obtained in this paper are indeed related. In this paper it is noted that the higher moments can be calculated using a technique due to Brockett, [2]. This calculation is also very difficult for higher dimensional systems. However the stability results of this paper do not depend on the higher moments.

The remainder of the paper is organized as follows. Section 2 provides the necessary background on convergence of random elements and Section 3 computes the related moments (mean and variance). Then, using the results developed in Sections 2 and 3, Section 4 through 9 explore the stability and convergence of the discrete system. Finally, Section 10 concludes the paper.

2. Background on the Convergence of Random Elements.

2.1. Random Variables.

First consider the case of random variables. Throughout, this subsection $(X_n :$

$n \geq 1$) denotes a sequence of random variables and X denotes a random variable, so that for each fixed $\omega \in \Omega$, $X(\omega), X_n(\omega) \in \mathbb{R}$; if necessary we can consider the extended real line $\mathbb{R} \bigcup \{-\infty, \infty\}$. The standard modes of convergence for random variables and their relationships are discussed in [1, 5]. Our focus will be *almost sure convergence*, also known as *convergence with probability one*. For completeness we recall the definition here. X_n converges almost surely (a.s.) to X, denoted $X_n \xrightarrow{a.s.} X$, if

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1.$$

A sufficient condition for almost sure convergence is found from the Borel-Cantelli Lemma as stated here.

PROPOSITION 2.1. *Assume the following holds for all* $\epsilon > 0$,

$$\sum_{n\geq 1} P\left(|X_n - X| \geq \epsilon\right) < \infty \tag{2.1}$$

Then $X_n \xrightarrow{a.s.} X$.

REMARK 2.2. *In general, condition* (2.1) *is a sufficient condition for almost sure convergence (in fact, it is sometimes referred to as complete convergence). But if* X_n *is an independent sequence of random variables then* (2.1) *is necessary and sufficient.*

The following result gives a standard way to apply Proposition 2.1 through moment calculations.

PROPOSITION 2.3. *If either*

1. $\sum_{n\geq 1} E\,|X_n| < \infty$

or

(2) $EX_n \to 0$ *and* $\sum_{n\geq 1} var\,(X_n) < \infty$.

then $X_n \xrightarrow{a.s.} X$.

Proof. We prove that (2) is a sufficient condition the proof for (1) is almost identical.

Let $Y_n = X_n - EX_n$. Fix $\epsilon > 0$, by Chebyshev's inequality we have,

$$\sum_{n\geq 1} P\left(|Y_n| > \epsilon\right) = \sum_{n\geq 1} P\left(|X_n - EX_n| > \epsilon\right) \leq \sum_{n\geq 1} \frac{var(X_n)}{\epsilon^2} < \infty,$$

by assumption. Therefore, by Borel-Cantelli, $Y_n \xrightarrow{a.s.} 0$. But

$$X_n = Y_n + EX_n,$$

and by assumption the constant sequence $EX_n \to 0$. Therefore $X_n \xrightarrow{a.s.} 0$. □

2.2. Extension to Random Vectors. These results are easily extended to random vectors because convergence almost surely for random vectors can be proven component-wise. Namely, for each n, let $\mathbf{X}_n$ and $\mathbf{X}$ be random vectors on $\mathbb{R}^m$. Then we have

$$\mathbf{X}_n \xrightarrow{a.s.} \mathbf{X} \quad \text{iff} \quad X_{nj} \xrightarrow{a.s.} X_j \quad \text{for } 1 \leq j \leq m,$$

3. Mean and Variance. In this section we think of switched linear systems as being of the form of a bilinear stochastic control system

$$x_{n+1} = (u_1 A_1 + \cdots + u_k A_k) x_n \tag{3.1}$$

where the u_i's are random variables with $u_i \in \{0, 1\}$, $\sum_{i=1}^{k} u_i = 1$ and $P(u_i = 1) = \frac{1}{k}$ and each $A_i \in GL(n, R)$. We note that it is not necessary for the probabilities to be uniformly distributed as $\frac{1}{k}$ but it simplifies the notation and is the most commonly used distribution. We let the u_i's take value in the set $\{0, 1\}$ so that they are identically distributed but are not independent. We calculate the the first two moments–the mean and covariance using a simple averaging process and then calculate the higher moments using a process which is due to Brockett in his study of Volterra series and nonlinear control, [2]. Brockett in [3] does this calculation in a slightly different setting for other distributions. The construction that we use appears for the mean in [8]. The context in that paper was switching between numerical methods to improve accuracy in the numerical solution of ordinary differential equations.

Let $S = \{A_i : \ i = 1, \cdots, k, \quad A_i \in Gl(n, R)\}$. Let

$$Y_m = X_m X_{m-1} \cdots X_0$$

where each X_i is a random variable taking values in S with $P(X_j = A_i) = \frac{1}{k}$.

Now let the system be defined as

$$x_{m+1} = (\delta_{1,m} A_1 + \delta_{2,m} A_2 + \cdots + \delta_{k,m} A_k) x_m$$

with the property that for each i $\delta_{i,m} \in \{0, 1\}$ and

$$\sum_{i=1}^{k} \delta_{i,m} = 1.$$

We then have that each particular sample path is of the form

$$x_m = Y_{m-1} x_0.$$

THEOREM 3.1. *Let*

$$S^m = \{Y_m : \text{ taken over all sample paths}\},$$

then the mean value of the $S'_m s$, E_m, is given recursively by

$$E_m = (\frac{1}{k} \sum_{i=1}^{k} A_i) E_{m-1}.$$

Proof. This proof follows the proof in [8]. We will calculate the mean of S^m. Let $S_i^m = \{Y_m \in S^m : \ X_m = A_i\}$. It is clear that S^m is the disjoint union of the S_i^m.

Let

$$E_m = \frac{1}{k^m} \sum_{Y_m \in S^m} Y_m.$$

Decomposing this sum we have

$$\begin{aligned} E_m &= \frac{1}{k^m} \sum_{Y_m \in S^m} Y_m \\ &= \frac{1}{k^m} \sum_{i=1}^{k} \sum_{Y_{m-1} \in S^{m-1}} A_i Y_{m-1} \\ &= \frac{1}{k} \sum_{i=1}^{k} \frac{1}{k^{m-1}} \sum_{Y_{m-1} \in S^{m-1}} A_i Y_{m-1} \\ &= \frac{1}{k} \sum_{i=1}^{k} A_i E_{m-1} \\ &= (\frac{1}{k} \sum_{i=1}^{k} A_i) E_{m-1} \end{aligned} \tag{3.2}$$

Thus we have the mean of the set S^m computed recursively. □

We now calculate the covariances.

THEOREM 3.2. *Let*

$$S^m = \{Y_m : \text{ taken over all sample path}\},$$

then the covariance of the $S'_m s$ *is given by*

$$V_m = C_m - E_m E'_m$$

and C_m *is generated recursively as*

$$C_{m+1} = \frac{1}{k} \sum_{i=1}^{k} A_i C_m A'_i.$$

Proof. We now calculate the covariance. Let

$$V_m = \frac{1}{k^m} \sum_{Y_m \in S^m} (Y_m - E_m)(Y_m - E_m)'.$$

Again we decompose the sum as

$$\begin{aligned} V_m &= \frac{1}{k^m} \sum_{Y_m \in S^m} (Y_m - E_m)(Y_m - E_m)' \\ &= \frac{1}{k} \sum_{i=1}^{k} \frac{1}{k^{m-1}} \sum_{Y_{m-1} \in S^{m-1}} (A_i Y_{m-1} - E_m) \times \\ &\qquad (A_i Y_{m-1} - E_m)' \end{aligned} \tag{3.3}$$

On the other hand we have

$$V_m = \frac{1}{k^m} \sum_{Y_m \in S^m} Y_m Y_m' - E_m E_m' \tag{3.4}$$

We now evaluate the sum

$$\begin{aligned}
&\frac{1}{k^{m-1}} \sum_{Y_{m-1} \in S^{m-1}} (A_i Y_{m-1} - E_m)(A_i Y_{m-1} - E_m)' = \\
&\quad \frac{1}{k^{m-1}} \sum_{Y_{m-1} \in S^{m-1}} [A_i Y_{m-1} Y_{m-1}' A_i' - \\
&\quad E_m Y_{m-1}' A_i' - A_i Y_{m-1} E_m' + E_m E_m'] \\
&= \frac{1}{k^{m-1}} \sum_{Y_{m-1} \in S^{m-1}} [A_i Y_{m-1} Y_{m-1}' A_i'] - E_m E_m'
\end{aligned}$$

Now define

$$C_m = \frac{1}{k^m} \sum_{Y_m \in S^m} Y_m Y_m'.$$

Thus we finally have

$$\begin{aligned}
V_m &= \frac{1}{k^m} \sum_{Y_m \in S^m} Y_m Y_m' - E_m E_m' \\
&= C_m - E_m E_m' \\
&= \frac{1}{k} \sum_{i=1}^{k} A_i C_{m-1} A_i' - E_m E_m' \qquad (3.5)
\end{aligned}$$

This completes the derivation. □

We then have two linear recurrences that determine the mean and covariance of the S^m under the assumption of uniform probability distribution.

$$E_{m+1} = (\frac{1}{k} \sum_{i=1}^{k} A_i) E_m \tag{3.6}$$

$$C_{m+1} = \frac{1}{k} \sum_{i=1}^{k} A_i C_m A_i' \tag{3.7}$$

Now let $x_{n+1} = A x_n$ where $x_n \in R^m$ and A is an $m \times m$ constant real matrix. Let

$$X_n^{[p]} = \begin{pmatrix} x_{1,n}^p \\ x_{1,n}^{p-1} x_{2,n} \\ \vdots \\ x_{m,n}^{p-1} x_{m-1,n} \\ x_{m,n}^p \end{pmatrix}$$

where this notation is intended to denote all monomials of degree p. We will show that there exists a real constant matrix, $A_{[p]}$ so that

$$X_{n+1}^{[p]} = A_{[p]} X_n^{[p]}.$$

Consider the monomial

$$x_{1,n+1}^{n_1} x_{2,n+1}^{n_2} \cdots x_{m,n+1}^{n_m}$$

where $n_1 + \cdots + n_m = p$ and each $n_i \geq 0$. Now

$$x_{i,n+1} = e_i' A x_n$$

and hence

$$x_{1,n+1}^{n_1} x_{2,n+1}^{n_2} \cdots x_{m,n+1}^{n_m} = (e_1' A x_n)^{n_1} \cdots (e_m' A x_n)^{n_m}$$

Each of the terms on the right is homogeneous linear and hence the product is the sum of monomials of degree p and thus one row of the matrix $A_{[p]}$ is determined. We show examples for $n = 2$ and $p = 2$ and $p = 3$.

$$\begin{pmatrix} x_1^2 \\ x_1 x_2 \\ x_2^2 \end{pmatrix}_{n+1} = \begin{pmatrix} a_{11}^2 & 2a_{11}a_{12} & a_{12}^2 \\ a_{11}a_{21} & a_{12}a_{21} + a_{11}a_{22} & a_{12}a_{22} \\ a_{21}^2 & 2a_{21}a_{22} & a_{22}^2 \end{pmatrix} \begin{pmatrix} x_1^2 \\ x_1 x_2 \\ x_2^2 \end{pmatrix}_n$$

$$\begin{pmatrix} x_1^3 \\ x_1^2 x_2 \\ x_1 x_2^2 \\ x_2^3 \end{pmatrix}_{n+1} =$$

$$\begin{pmatrix} a_{11}^3 & 3a_{11}^2 a_{12} & 3a_{11}a_{12}^2 & a_{12}^3 \\ a_{11}^2 a_{21} & a_{11}^2 a_{22} + 2a_{11}a_{21}a_{12} & a_{12}^2 a_{21} + 2a_{11}a_{12}a_{22} & a_{12}^2 a_{22} \\ a_{11}a_{21}^2 & a_{12}a_{21}^2 + 2a_{11}a_{21}a_{22} & a_{11}a_{22}^2 + 2a_{12}a_{21}a_{22} & a_{12}a_{22}^2 \\ a_{21}^3 & 3a_{21}^2 a_{22} & 3a_{21}a_{22}^2 & a_{22}^3 \end{pmatrix} \times$$

$$\begin{pmatrix} x_1^3 \\ x_1^2 x_2 \\ x_1 x_2^2 \\ x_2^3 \end{pmatrix}_n$$

We consider the monomials for the system 3.1. Note that

$$\delta_i \delta_j = \begin{cases} 0 & i \neq j \\ \delta_i & i = j \end{cases}$$

and hence if we construct the system as above we have

$$X_{n+1}^{[p]} = (\delta_1 A_{[p],1} + \cdots + \delta_k A_{[p],k}) X_n^{[p]}. \tag{3.8}$$

we thus have the following theorem as a direct application of theorem 3.1.

THEOREM 3.3. *Let*

$$X_{n+1}^{[p]} = (\delta_1 A_{[p],1} + \cdots + \delta_k A_{[p],k}) X_n^{[p]}$$

then the expected value of the p−moments is given by

$$E(X^{[p]})_{n+1} = \frac{A_{[p],1} + \cdots + A_{[p],k}}{k} E(X^{[p]})_n.$$

4. Stability of system. In this section we will prove the following theorem.

THEOREM 4.1. *If the system*

$$x_{m+1} = (\delta_{1,m} A_1 + \delta_{2,m} A_2 + \cdots + \delta_{k,m} A_k) x_m$$

is stable for all choices of the $\delta_{i,m}$, $\delta_{i,m} \in \{0,1\}$ $\sum_i \delta_{i,m} = 1$, *then*

$$E_{m+1} = (\frac{1}{k} \sum_{i=1}^{k} A_i) E_m$$

$$C_{m+1} = \frac{1}{k} \sum_{i=1}^{k} A_i C_m A_i'$$

are both stable.

Proof. We construct a tree from a natural partial ordering on the switching sequences and to prove this theorem we will make essential use of König's finitary tree theorem, [11].

THEOREM 4.2 (König). *Every infinite finitary tree has an infinite branch.*

We define an ordering on the set of all finite sequences of numbers 1 through k. Let $\gamma^n = (\gamma_n, \gamma_{n-1}, \cdots, \gamma_1)$ where $\gamma_i \in \{0, 1, \cdots, k\}$. Let

$$S_n = \{\gamma_n : \text{ over all choices of } \gamma_i\}.$$

We let S_0 be the set consisting of the empty sequence which we will denote by $\emptyset$. We will say that an element of S_n has height n. Let $S = \cup_n S_n$. Let $x, y \in S$ and suppose that the height of x is less than the height of y. We define $x \prec y$ if and only if $x = (\gamma_n, \cdots, \gamma_1)$ and $y = (\beta_k, \cdots, \gamma_1, \gamma_n, \cdots, \gamma_1)$. Note that $\emptyset \prec x$ for all x of positive height. Note that S_n is finite and has exactly k^n elements. So $(S, \prec)$ is a rooted tree and since the number of elements of height n is finite it is a finitary tree.

Now let $\{\delta_{im} : \ m = 1, 2, \cdots\}$ be any infinite sequence of 1's with corresponding matrices $Y_n = (\delta_{1,n} A_1 + \delta_{2,n} A_2 + \cdots + \delta_{k,n} A_k)(\delta_{1,n-1} A_1 + \delta_{2,n-1} A_2 + \cdots + \delta_{k,n-1} A_k) \cdots (\delta_{1,1} A_1 + \delta_{2,1} A_2 + \cdots + \delta_{k,1} A_k)$. We now assume that the systems is

stable for all choices of switching sequences. Now given any fixed epsilon, $\epsilon > 0$ here exists a N such that $\|Y_n\| < \epsilon$ for $n > N$. Let N be the smallest N that works. Now for each m there exists a unique $\delta_{i(m),m} = 1$ and we define a finite sequence $(i(N), i(N-1), \cdots, i(1)) \in S_N$. Then for a fixed ϵ there is a mapping from the set of all switching sequences into S. Let the image of the set of all sequences be denoted by R. Let $\hat{R}$ be the smallest rooted tree that contains S. Thus $\hat{R}$ is a finitary tree and hence if it is infinite then it contains an infinite branch. This contradicts the fact that the system is stable and therefore the tree must be finite. Thus there exists an N_0 such so that for all Y_n, $n > N_0$ implies that $\|Y_n\| < \epsilon$.

Now let $n > N_0$ and calculate the average value of S_n.

$$\begin{aligned} \|\frac{1}{k^n} \sum_{Y_n \in S_n} Y_n\| &\leq \frac{1}{k^n} \sum_{Y_n \in S_n} \|Y_n\| \\ &< \frac{1}{k^n} \sum_{Y_n \in S_n} \epsilon \\ &= \epsilon \end{aligned}$$

Hence, for all $n > N_0$ $\|E_n\| < \epsilon$. The calculation for the covariance is similar to the construction of the mean which concludes the proof.

5. Almost Sure Convergence. In this section we show that a necessary and sufficient condition for almost sure convergence of the system 3.1 is that the matrix

$$\frac{A_1 + \cdots + A_k}{k}$$

has all of its eigenvalues interior to the unit circle.

THEOREM 5.1. *The system*

$$x_{n+1} = (\delta_1 A_1 + \cdots + \delta_k A_k) x_n$$

converges to zero almost surely if and only if every eigenvalue of the matrix

$$\frac{A_1 + \cdots + A_k}{k}$$

lies interior to the unit circle.

Proof. The proof is a direct consequence of Theorem 3.1 and Proposition 2.3. We need only show that

$$\sum_n Ex_n < \infty.$$

So suppose that the matrix has its eigenvalues interior to the unit circle. Then by the recurrence we have

$$Ex_n = \left(\frac{A_1 + \cdots + A_k}{k}\right)^n Ex_0$$

and there exist matrices $F_1, \cdots, F_k$ such that

$$Ex_n = (\lambda_1^n F_1 + \cdots + \lambda_k^n F_k) Ex_0$$

and so the sum can be written as

$$\begin{aligned} \sum_n Ex_n &= \sum_n \lambda_1^n F_1 + \cdots + \sum_n \lambda_k^n F_k) Ex_0 \\ &= (\frac{1}{1-\lambda_1} F_1 + \cdots + \frac{1}{1-\lambda_k} F_k) Ex_0 \\ &< \infty \end{aligned}$$

since the expected value of the vector converges so does each component. also note that if any eigenvalue is outside the unit circle then the geometric sum with that eigenvalue diverges to infinity. For an eigenvalue λ on the unit circle then λ^n does not converge to zero and hence the sum cannot converge and further more Ex_n does not converge to zero.

6. Counter examples. One is tempted to conjecture that a necessary and sufficient condition for stability is that the mean and covariance are stable. In this section we show that this not true by producing an example for $k = 2$ of a system for which the mean and covariance are stable (but with eigenvalues very close to 1) and there exists a switching sequence that renders the system unstable. In view of the failure it is tempting to conjecture that if all of the moments converge to zero then the system is stable for all switching sequences. We produce an example for which all of them moments converge to zero but for which there are sequences that drive the system away from zero. This reminiscent of the example in [6] of a system which is stable but for which there is no quadratic Lyapunov function.

We will now give a counterexample to show that the system

$$x_{n+1} = (\delta_n A_1 + (1 - \delta_n) A_2) x_n$$

is not always stable for every sequence of $\{\delta_n\}$ even though the mean and covariance are stable. Let

$$A_1 = \begin{pmatrix} 0.9739 & 0.0098 \\ -0.9772 & 0.9739 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 0.9719 & 0.0975 \\ -0.0975 & 0.9719 \end{pmatrix}$$

The two matrices satisfy the conditions that:

1) A_1^{-1} and A_2^{-1} exist.
2) As $k \to \infty$, $A_1^k \to 0$ and $A_2^k \to 0$.

Then the average of A_1 and A_2 is:

$$\frac{A_1 + A_2}{2} = \begin{pmatrix} 0.9729 & 0.0536 \\ -0.5374 & 0.9729 \end{pmatrix}$$

The corresponding eigenvalues for the average are $\lambda_1 = 0.9729 + 0.1698i$, $\lambda_2 = 0.9729 - 0.1698i$, with absolute value of the eigenvalues $0.9876 \leq 1$.

Writing the covariance dynamics as a matrix C we have

$$C = \begin{pmatrix} 0.9465 & -1.0535 & 0.4838 \\ 0.0525 & 0.9370 & -0.5268 \\ 0.0048 & 0.1050 & 0.9465 \end{pmatrix}$$

The eigenvalues of C are $\lambda_1 = 0.9948$, $\lambda_2 = 0.9176 + 0.3320i$, and $\lambda_3 = 0.9176 - 0.3320i$, with maximum absolute value of eigenvalue $0.9948 \leq 1$. Since the maximum eigenvalues of E and C are less than 1 both E_n and C_n are asymptotic stable. If for all choice of δ_n, the system

$$x_{n+1} = (\delta_n A_1 + (1 - \delta_n)A_2)x_n$$

is always stable, we are expecting that the switching curve goes eventually to 0. However, in this example, we switch between the two systems by the following manner: We start with system A_1, and switch the system to A_2 when arriving at the furthest point on the flow of A_1. Then we switch the system back to A_1 while it arrive at the furthest point on the flow of system A_2. Continuing with this switching method, we are able to drive the system away from the origin.The figure below shows the trajectory of the switched system.It is possible to calculate the exact sequence of switches that drive the system to infinity.

Now consider the system

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{n+1} = \left(\delta_n \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + (1 - \delta_n) \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_n$$

For this system there are exactly two switching sequences that do not result in $x_n = 0$ for some finite n, namely the sequences $\delta_n = \frac{1+(-1)^n}{2}$ and $\delta_n = \frac{1+(-1)^{n+1}}{2}$. Note that

$$\frac{A_1 + A_2}{2} = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and the eigenvalues of this matrix are $\pm\frac{1}{2}$ and hence the means converge to 0. The second moments are determined by the eigenvalues of the matrix

$$\frac{1}{2} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

and the eigenvalues of this matrix are $\pm\frac{1}{2}$ and 0. In general the p^{th} moments are determined by a $p+1 \times p+1$ matrix that has a 1 in the upper right hand corner and a 1 in the lower left hand corner and zeros elsewhere. The eigenvalues of such a matrix are $\pm\frac{1}{2}$ and 0 with a multiplicity of p-1. Thus all of the moments converge to zero but there are two sequences that do not converge to zero.

7. Lyapunov Functions. Much of the stability analysis of switching systems has centered around the existence of Lyapunov functions. In this section, we construct a sequence of rational Lyapunov functions of increasing degree. Unfortunately, the examples of the previous section demonstrate that even when focusing on this countable set of Lyapunov functions there is still no guarantee that the system will be stable for all possible switching sequences.

An elementary result for linear systems theory is that the system

$$x_{n+1} = Ax_n \tag{7.1}$$

is stable if and only if the spectrum of A is interior to the unit disk. Then, using a conformal mapping, the system is stable if and only if the continuous time system

$$\dot{x} = (A-I)(A+I)^{-1}x$$

is stable. We know that if this system is stable, then there exists a quadratic Lyapunov function, with positive definite matrix P so that

$$[(A-I)(A+I)^{-1}]'P + P[(A-I)(A+I)^{-1}] = -I.$$

We thus have the following theorem.

THEOREM 7.1. *Consider the system*

$$x_{n+1} = (\delta_n A + (1-\delta_n)B)x_n.$$

The systems

$$EX^{[p]}_{n+1} = (\delta_n A_{[p]} + (1-\delta_n)B_{[p]})X^{[p]}_n$$

are stable if and only if there exists a sequence of positive definite matrices Q_p such that

$$[(A_{[p]} + B_{[p]} - 2I)(A_{[p]} + B_{[p]} + 2I)^{-1}]'Q_p +$$

$$Q_p[(A_{[p]} + B_{[p]} - 2I)'(A_{[p]} + B'_{[p]} + 2I)^{-1}]' = -I.$$

The proof is immediate.

Because the matrices $A_{[p]}$ are polynomial in the entries of A, the theorem results in a sequence of higher order Lyapunov functions for the system

$$x_{n+1} = (\delta_n A + (1-\delta_n)B)x_n.$$

However, the second example above shows that even if this sequence of Lyapunov functions exists, there is still no guarantee of the convergence of all switching sequences. This is reminiscent of the result in [6], which shows that there exist stable continuous time switching systems for which there is no quadratic Lyapunov function.

8. Stability of E_n and C_n. We now assume that E_n and C_n are asymptotically stable. We begin with a series of lemmas.

LEMMA 8.1. *The matrix C_n is positive definite.*

Proof. First note that $C_0 = I$. Assuming C_n is positive definite we have $x'C_{n+1}x = \frac{1}{k}\sum_{i=1}^{k} x'A_iC_mA_i'x > 0$ and the lemma follows by induction.

We are assuming that both the average and the covariance are stable and since they are generated by a linear recurrence this implies that they are exponentially stable. Thus we have that for some suitable norm $\|C_{n+1}\| < \|C_n\|$. The adjoint of C_n, C_n^* is generated by the recurrence

$$C_{n+1}^* = \frac{1}{k}\sum_{i=1}^{k} A_i'C_m^*A_i.$$

It is more convenient to work with the adjoint than the covariance directly. We state the following as a lemma but it is obvious.

LEMMA 8.2. *Let Z be any $n \times n$ matrix.*

$$\|Z'C_{n+1}^*Z\| \leq \|Z'C_n^*Z\|.$$

Let $Y_n \in S_n$ and calculate

$$\begin{aligned} Y_n'C_{m+1}^*Y_n &= \frac{1}{2^k}\sum_{i=1}^{k} Y_n'A_i'C_mA_iY_n \\ &= \frac{1}{2^k}\sum_{i=1}^{k} Y_n^{i\prime}C_mY_n^i \end{aligned}$$

Where the average is taken over all immediate successors of Y_n. Thus we have the important lemma.

LEMMA 8.3. *For all $x_0 \in \mathbb{R}^n$*

$$x_o'Y_n'C_m^*Y_nx_0 \geq \frac{1}{2^k}\sum_{i=1}^{k} x_0'Y_n^{i\prime}C_m^*Y_n^ix_0$$

From the lemma we see that the system is "on the average" stable. However there may be a sequence

$$Y_0 \prec Y_1 \prec \cdots$$

for which for every n and k

$$Y_{n+1}'C_kY_{n+1} > Y_n'C_kY_n.$$

Our goal is to show that this cannot happen for a "large" set of switching sequences.

9. Almost every sequence. Let δ be any infinite sequence $\delta = \{\delta_{n,i}\}_{i=0}^{k}$ as in the definition of the system. We define an integer

$$\delta_n = 0\delta_{n,0} + 1\delta_{n,1} + 2\delta_{n,2} + \cdots + (k-1)\delta_{n,k-1}$$

and using this we define a real number as

$$r(\delta) = \sum_{n=1}^{\infty} \delta_n k^{-n}.$$

Note that δ_n is an integer between 0 and $k-1$. This extends to a map from the formal sequence $\{\sum_{i=0}^{k-1} \delta_{n,i} A_i\}$ to $\mathbb{R}$. We now prove the following theorem. Given an infinite sequence $\{\delta\}$ let $Y_n = X_n \cdots X_1$ where $X_i = \sum_{j=0}^{k-1} \delta_{n,j} A_j$.

LEMMA 9.1. *Let*

$$S = \{\{\delta\} : \lim_{n\to\infty} Y_n \text{ does not converge to } 0\}.$$

If E_n converges to 0 then $r(S)$ does not contain any non empty open interval.

Proof. Suppose $r(S)$ contains an open interval. Then for some k the interval $[a + k^{-j-1}, a + k^{-j}]$ is contained in $r(S)$ where $a = \sum_{i=1}^{k-2} \delta_i k^{-i}$. Then every number of the form $a + \sum_{i=k}^{\infty} \delta_i 2^{-i}$ is in the interval for every choice of δ_i. Thus in S_n there are k^{n-j} elements Y_n with norm greater than ϵ. Thus

$$E_n = \frac{1}{k^n} \sum Y_n > \frac{k^{n-j}}{k^n} \epsilon = \frac{\epsilon}{k^j}$$

Since j is fixed E_n is bounded away from 0 and hence does not converge.

We now state an important conjecture. This conjecture is very much in the spirit of the results in [9] and [10]. The convergence results of those papers are all of the form convergence with probability 1. That is, there is a possibly a set of Lebesgue measure 0 for which there is no convergence. For the counterexample we have not calculated the set of all destabilizing sequences but it is clear that they must have a very special form that leads us to believe that they form a set of measure 0.

THEOREM 9.2 (conjecture). *Let*

$$S = \{\{\delta\} : \lim_{n\to\infty} Y_n \text{ does not converge to } 0\}.$$

If E_n converges to 0 then $r(S)$ does not contain any set of positive Lebesgue measure.
Some thoughts on a proof: Assume $r(S)$ has positive Lebesgue measure, $\mu(r(S)) > 0$. From the definition of Lebesgue measure we have that

$$\mu(r(S)) = \inf \sum_n \mu(I_n)$$

$\cup I_n \supseteq S$. Thus we have that there exists n such that $\mu(I_n \cap S) > 0$. As in the proof of the lemma there exists an open interval of the form $I = (a + k^{-j-1}, a + k^{-j})$ in I_n and further more there must exist such an interval with the property that

$$\mu(I \cap S) > 0.$$

The idea of the proof would be to show that there are sufficiently many sequences in this set so that the expected value is bounded away from 0.

10. Conclusion. The theory of random products has been an important topic in statistics and mathematical physics for the last half century. It is easy to see the connection with the theory of switched linear systems. The two areas do not have identical interests. In statistics and in mathematical physics much of the emphasis has been and is on the eigenvalues of the products. This particular line has not been of interest in the theory of switched systems. In systems theory the ideas of stability and controllability along with ideas of how to approximate switched systems with more easily studied systems, [20], have been the main directions. One contribution of this paper is an attempt to use ideas from the two areas.

REFERENCES

[1] P. Brémaud, *An introduction to probabilistic modeling. Undergraduate Texts in Mathematics*, Springer-Verlag, New York, 1988.

[2] R. Brockett, *Volterra series and geometric control theory*, Automatica—J. IFAC 12:2(1976), pp. 167–176.

[3] R. Brockett, *Stochastic Control*, lecture notes, 2008.

[4] O. Costa, M. Fragoso, and R. Marquez, *Discrete-time Markov jump linear systems*, London: Springer; 2005.

[5] Y.S. Chow and H. Teicher, *Probability theory. Independence, interchangeability, martingales*, third edition, Springer Texts in Statistics. Springer-Verlag, New York, 1997.

[6] W. P. Dayawansa and C. Martin, *A converse Lyapunov theorem for a class of dynamical systems which undergo switching*, IEEE Trans. Automat. Control 44:4(1999), pp. 751–760.

[7] X. Dai, Y. Huang and M. Xiao, *Almost sure stability of discrete-time switched linear systems: a topological point of view*, SIAM J. Control Optim. 47(2008), pp. 2137–2156.

[8] M.P.B. Ekanayake, B. He, L. Huo, K. Kaphle, and C. Martin, *Control theory and the numerical solution of ODEs*, Communications in Informations and Systems, 8:3(2008), pp. 201–222.

[9] H. Furstenberg and H. Kesten, *Products of random matrices*, Ann. Math. Statist 31(1960), pp. 457–469.

[10] H. Furstenberg, *Noncommuting random products*, Trans. Amer. Math. Soc. 108(1963), pp. 377–428.

[11] D. König, *Theorie der Endlichen und Unendlichen Graphen: Kombinatorische Topologie der Streckenkomplexe*, Leipzig: Akad. Verlag. (1936).

[12] D. Liberzon, *Switching in systems and control. Systems & Control: Foundations & Applications*, Birkhuser Boston, Inc., Boston, MA, 2003.

[13] H. Lin and P.J. Antsaklis, *Stability and persistent distrubance attenuation properties for networked control systems: switched systems approach*, Inter. J. control, 78(2005), pp. 1447–1458.

[14] H. Lin and P.J. Antsaklis, *Stability and stabilizability of switched linear systems: a survey of recent results*, IEEE TAC, 54(2009), pp. 308–322.

[15] C. Martin, *Unique factorization of arithmetic functions*, Aequationes Math. 7 (1971), 211.

[16] S. A. Murphy, *Optimal dynamic treatment regimes*, J. R. Stat. Soc. Ser. B Stat. Methodol. 65:2(2003), pp. 331–366.

[17] S. A. Murphy, M.J. van der Laan, and J.M. Robins, *Marginal mean models for dynamic regimes*, J. Amer. Statist. Assoc. 96:456(2001), pp. 1410–1423.

[18] S. A. Murphy and D. Bingham, *Screeing experiments for Developing Dynamic Treatment Regimes*, Journal of the ASA, Vol 104, pp. 391–408.

[19] M. Ogura and C. Martin, *Linear switching systems and random products of matrices*, to appear MTNS2012.

[20] P.C. Perera and W.P. Dayawansa, *Approximation of a switched linear system by a non-switched homogeneous polynomial system*, J. Control Theory Appl. 4:1(2006), pp. 3–10.

[21] J. N. Tsitsiklis and V. D. Blondel, *The Lyapunov exponent and joint spectral radius of pairs of matrices are hard – when not impossible – to compute and to approximate*, Mathematics of Control, Signals, and Systems, 10:1(1997), pp. 31–40.

[22] N. Wang, *Switched systems as models of dynamic treatment scheduling*, Master's thesis, Texas Tech University, 2009.

[23] N. Wang, M. Egerstedt, and C. Martin, *Stability of switched linear systems and the convergence of random products*, Proceedings of the 48th CDC, 2009.