Stability of switched linear systems and the convergence of random products

N. Wang*, M. Egerstedt** and C. Martin*

Abstract— **In this paper we give conditions that a discrete time switched linear systems must satisfy if it is stable. We do this by calculating the mean and covariance of the set of matrices obtained by using all possible switches. 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 exists in the statistics literature. A nice byproduct of this research is that we use König's theorem of finatary trees. This may be the first use of this theorem in systems and control.**

I. 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, [10], gives a rather complete overview of the field. the main theorem cited there on stability for discrete time switched linear systems, (their Theorem 6, [9]), 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" switching sequences. an excellent and readable source for the general theory of switched linear systems is the monograph by Liberzon, [8]. 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, [5], [6], 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.

*Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX, 79409-1024, USA, **School of Electrical Engineering, Georgia Institute of Technology, Atlanta, Georgia

The work in this paper was motivated by a problem in area of dynamic clinical trials, [12], [13]. A characteristic of treating a terminal disease with multiple drug therapies is the timing of the drug treatments, e.g. [3], [12], [15]. The model is developed in the following manner: We assume that some measure of quality, x, is being measured and that it is scaled so that 0 represents death and 1 represents remission. For each treatment we assume a model of the form $x_{n+2} = a_i x_{n+1} + b_i x_n + \epsilon_{n,i}$, where the parameters a_i, b_i are to be determined based on the treatment and $\epsilon_{n,i}$ are iid (independent and identically distributed). Then, for two treatments, we have the combined model

$$\begin{aligned} x_{n+2} &= \delta_n(a_1 x_{n+1} + b_1 x_n + \epsilon_{n,1}) \\ &\quad +(1-\delta_n)(a_2 x_{n+1} + b_2 x_n + \epsilon_{n,2}) \end{aligned}$$

where $\delta_n \in \{0,1\}$. We consider δ_n to be a control variable that represents the treatment to be applied. We show in [3] and [15] that the best possible switching times occur when the system is driven by one treatment as far from zero as possible. At this point the patient should be switched to the second treatment. This is in contrast to the usual practice of waiting until it is clear that the treatment is failing before applying the second treatment.

This work led naturally to the study of the stability of discrete time switched systems. 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 exists in the statistics literature. A nice byproduct of this research is that we use König's theorem of finatary trees, [11]. This may be the first use of this theorem in systems and control.

II. Mean and Variance

In this section we think of the switched linear system as being of the form of a bilinear stochastic control system

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

where the u_is are random variables with $u_i \in \{0, 1\}$, $\sum_{i=1}^{k} u_i = 1$ and $P(\mu_i = 1) = \frac{1}{k}$ 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_is 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. Brockett in [1] does this calculation in a slightly different setting for other distributions. The construction that we use appears for the mean in [4]. 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 2.1: 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 [4]. 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

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

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

We now calculate the covariances.

Theorem 2.2: 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

$$\begin{aligned}
V_m &= \frac{1}{k^m} \sum_{Y_m \in S^m} (Y_m - E_m)(Y_m - E_m)' \\
&= \frac{1}{k} \sum_{i=1}^{k} \frac{1}{k^{m-1}} \sum_{Y_{m-1} \in S^{m-1}} (A_i Y_{m-1} - E_m) \times \\
&\quad (A_i Y_{m-1} - E_m)' \qquad (2)
\end{aligned}$$

On the other hand we have

$$V_m = \frac{1}{k^m} \sum_{Y_m \in S^m} Y_m Y'_m - E_m E'_m \qquad (3)$$

We now evaluate the sum

$$\frac{1}{k^{m-1}}\sum_{Y_{m-1}\in S^{m-1}}(A_iY_{m-1}-E_m)(A_iY_{m-1}-E_m)' =$$
$$\frac{1}{k^{m-1}}\sum_{Y_{m-1}\in S^{m-1}}[A_iY_{m-1}Y'_{m-1}A'_i -$$
$$E_mY'_{m-1}A'_i - A_iY_{m-1}E'_m + E_mE'_m]$$
$$= \frac{1}{k^{m-1}}\sum_{Y_{m-1}\in S^{m-1}}[A_iY_{m-1}Y'_{m-1}A'_i] - E_mE'_m$$

Now define

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

Thus we finally have

$$\begin{aligned} V_m &= \frac{1}{k^m}\sum_{Y_m\in S^m}Y_mY'_m - E_mE'_m \\ &= C_m - E_mE'_m \\ &= \frac{1}{k}\sum_{i=1}^{k}A_iC_{m-1}A'_i - E_mE'_m \end{aligned} \quad (4)$$

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.

$$E_{m+1} = (\frac{1}{k}\sum_{i=1}^{k}A_i)E_m \quad (5)$$

$$C_{m+1} = \frac{1}{k}\sum_{i=1}^{k}A_iC_mA'_i \quad (6)$$

III. Stability of System

In this section we will prove the following theorem.

Theorem 3.1: 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

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

$$C_{m+1} = \frac{1}{k}\sum_{i=1}^{k}A_iC_mA'_i$$

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 finatary tree theorem, [7].

Theorem 3.2 (König): Every infinite finatary 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 finatary 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 finatary 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.

One is tempted to conjecture that a necessary and sufficient condition for stability is that the mean and covariance are stable. In the next 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. This reminiscent of the example in [2] of a system which is stable but for which there is no quadratic Lyapunov function.

IV. Counterexample

In this section, we will 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.

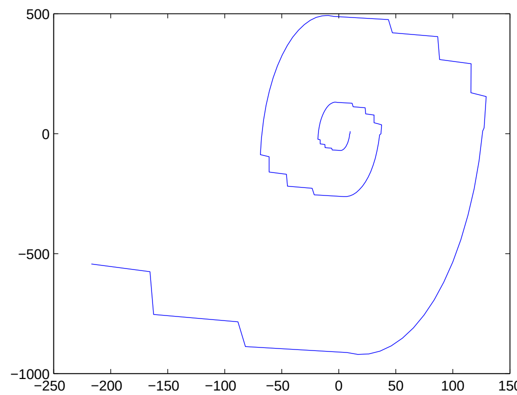

Fig. 1. Trajectory of the Switching system

V. 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 5.1: 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 5.2: 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 5.3: For all $x_0 \in \mathbb{R}^n$

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

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

$$Y_{n+1}' C_k Y_{n+1} > Y_n' C_k Y_n.$$

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

VI. 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 6.1: 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 [5] and [6]. The convergence results of those papers are all of the form convergence with probability 1. That is, there is a possibly a set of Lebesque 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 6.2 (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 Lebesque measure.

Some thoughts on a proof: Assume $r(S)$ has positive Lebesque measure, $\mu(r(S)) > 0$. From the definition of Lebesque 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.

VII. Conclusion

The theory of random products as 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, [14], have been the main directions. One contribution of this paper is an attempt to use ideas from the two areas.

References

[1] R. Brockett, Stochastic Control, lecture notes, 2008.

[2] W. P. Dayawansa, C. Martin, A converse Lyapunov theorem for a class of dynamical systems which undergo switching. IEEE Trans. Automat. Control 44 (1999), no. 4, 751–760.

[3] Y. Du, Simulation of the optimal timing of a multiple drug regime, Master's thesis, Texas Tech University, 2009

[4] M.P.B. Ekanayake, B. He, L. Huo, K. Kaphle and C. Martin, Control theory and the numerical solution of ODEs, to appear, Communications in Informations and Systems

[5] H. Furstenberg, H. Kesten, Products of random matrices. Ann. Math. Statist 31 1960 457–469.

[6] H. Furstenberg, Noncommuting random products. Trans. Amer. Math. Soc. 108 1963 377–428

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

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

[9] H. Lin, P.J. Antsaklis, Stability and persistent distrubance attenuation properties for networked control systems: switched systems approach, Inter. J. control, 78, 1447-1458, 2005.

[10] H. Lin, P.J. Antsaklis, Stability and stabilizability of switched linear systems: a survey of recnet results, IEEE TAC, 54, 2009, 308-322.

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

[12] S. A. Murphy, Optimal dynamic treatment regimes. J. R. Stat. Soc. Ser. B Stat. Methodol. 65 (2003), no. 2, 331–366.

[13] S. A. Murphy, M.J. van der Laan, J.M. Robins, Marginal mean models for dynamic regimes. J. Amer. Statist. Assoc. 96 (2001), no. 456, 1410–1423.

[14] P.C. Perera, W.P. Dayawansa, Approximation of a switched linear system by a nonswitched homogeneous polynomial system. J. Control Theory Appl. 4 (2006), no. 1, 3–10.

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