Estimation entropy for Regular Linear Switched Systems

G. S. Vicinansa and D. Liberzon

February 11, 2020

Abstract

In this paper, we introduce the concept of regular switched systems and we present some consequences of this property on the estimation entropy's calculation. For systems of such a class, one can derive a formula for the estimation entropy in terms of the system's Lyapunov exponents. We also present some sufficient conditions on the system that guarantee regularity. Some of those conditions include the cases of periodic switching, simultaneously triangularizable systems, and a large class of randomly switched systems that contains Markov Jump Linear Systems (MJLS) as a special case. For that last part, we use tools from ergodic theory to draw conclusions that hold almost surely.

1 Introduction

In recent years a lot of attention has been placed on the problem of estimation of systems subject to limited data-rate [8, 6]. Most of the attention those systems have received is due to the ever growing number of examples of systems that possess limited information flow between sensors and actuators.

A question that arises naturally is related to the minimum data rate necessary to accomplish a desired task, such as control or estimation. Normally, these questions have some type of entropy notion as an answer [9]. Trying to solve the case of estimation with minimum data-rate, the concept of estimation entropy was introduced in [6]. The goal in that paper was determining the minimum average data-rate necessary to estimate the state with an error that decreases exponentialy fast, with a prescribed exponential rate $\alpha \geq 0$. Nonetheless, it is not easy to calculate the value of the estimation entropy, with the notable exception of linear time invariant systems.

A class of systems for which the entropy concepts have recently attracted attention is that of switched systems [3]. Further, some recent works present inequalities for the topological entropy of linear switched systems [13] and the estimation entropy of nonlinear switched systems [10]. However, almost all results known are bounds that are sometimes very loose. In the present paper, we derive a formula for the estimation entropy of linear switched systems under mild restrictions on the switching.

First, we present a class of linear switched systems that are called regular systems. We show that for this class of systems, under some other technical assumptions, it is possible to write down a formula for the estimation entropy in terms of the system's Lyapunov exponents. This work can be seen as a specialization of the work [13] where general linear switched systems were studied, although most of the results presented there were bounds for the topological entropy, with no closed expressions for most of the matricial cases. Also, this present work extends the class of systems for which we know how to calculate the estimation entropy, which was restricted essentially to the linear time invariant case.

In Section 2 we present the preliminary results and the definition of regular linear time varying system in terms of the Lyapunov exponents of a system. Afterwards, in Section 3 we prove the formula for the estimation entropy for the class of discrete linear time varying (LTV) systems. In the sequel, in Section 4 we present sufficient conditions for regularity, most of those conditions are related to important classes of systems such as periodicaly switched linear systems and randomly switched linear systems, which includes Markov Jump Linear Systems. Finally, Section 5 concludes the paper and presents future research directions.

Notations: We denote by $||\cdot||$ a norm in a finite dimensional vector space, unless specified otherwise it can be taken to be any norm. Let $\mathbb{R} = (-\infty, \infty)$, $\mathbb{Z}_{\geq 0} = \{0, 1, \cdots\}$ the nonegative integers, $\mathbb{N} = \{1, 2, \cdots\}$ the set of natural numbers. For any set E, we denote by $\#E$ its cardinality. For subsets of $\mathbb{R}^d$ we denote $\text{vol}(E)$ the volume of the set (its Lebesgue measure). We also denote by $\dim(V)$ the dimension of a linear vector space V. Also, for any $x > 0$, $\log x$ is the logarithm with base e.

Given a matrix sequence $(A_n)_{n\in\mathbb{N}}$, we denote the product $A^{(n)} = A_n \cdots A_2 A_2$. We denote by $\text{Gl}(d, \mathbb{R})$ the general linear group of d by d matrices over the real field, and $\mathcal{M}(d, \mathbb{R})$ the set of all matrices over the reals. We denote $\det(A)$ and $\text{Tr}(A)$ the determinant and the trace of the matrix A, respectively. Further, $I_d \in \text{Gl}(d, \mathbb{R})$ is the identity matrix. Additionally, denote by $\mathcal{P}_B\left(\{v_i\}_{i=1}^k\right)$ the parallelepiped defined by $\{\alpha_i B v_i : \alpha_i \in [0, 1]\}$, where $\{v_i\}_{i=1}^k \subset \mathbb{R}^d$ is a set of vectors and $B \in \mathcal{M}(d, \mathbb{R})$. Furthermore, $\text{vol}\left(\mathcal{P}_B\left(\{v_i\}_{i=1}^k\right)\right)$ is the volume of the parallelepiped given by $\sqrt{\det\left((BV)^*(BV)\right)}$, where V is the matrix with columns v_i.

2 Preliminaries and Definitions

In this section we present some definitions and preliminary results needed for the discussion that follows. First we recall the definition of the estimation entropy [6] for a dynamic system given by the following equation:

$$\dot{x} = f(x), \quad x(0) \in K, \tag{1}$$

where K is the set of initial conditions that is compact and $K^o \neq \emptyset$. We denote by $\xi(x, t)$ the solution to the initial value problem for the initial condition $x(0) = x$ and time $t \in \mathbb{R}$.

Definition 2.1. *Let $\alpha \geq 0$ and $T \geq 0$ be the time horizon. For every $\epsilon > 0$, we call a finite set of functions $\hat{X} = \{\hat{x}_1(\cdot), \cdots, \hat{x}_N(\cdot)\}$, from $[0, T]$ to $\mathbb{R}^d$, a (T, ϵ, α, K)-approximating set if for every $x \in K$ initial condition, there exists $\hat{x}_i \in \hat{X}$ such that $||\xi(x, t) - \hat{x}_i(t)|| < \epsilon e^{-\alpha t}, \quad \forall t \in [0, T]$.*

Let $s_{est}(T, \epsilon, \alpha, K)$ be the minimum cardinality of a (T, ϵ, α, K)- approximating set. We define the estimation entropy as

$$h_{est}(\alpha, K) = \lim_{\epsilon \to 0} \limsup_{T \to \infty} \frac{1}{T} \log s_{est}(T, \epsilon, \alpha, K)$$

Another useful definition is that of $(T, \epsilon, \alpha, K) -$ *spanning set*

Definition 2.2. *Let $\alpha \geq 0$ and $T \geq 0$ be the time horizon. For every $\epsilon > 0$, we call a finite set of points $S = \{x_1, \cdots, x_N\} \subset K$ a (T, ϵ, α, K)- spanning set if for every $x \in K$ initial state, there exists $x_i \in S$ such that $||\xi(x, t) - \xi(x_i, t)|| < \epsilon e^{-\alpha t}, \quad \forall t \in [0, T]$.*

Let $s^*_{est}(T, \epsilon, \alpha, K)$ be the minimum cardinality of a (T, ϵ, α, K)- spanning set. We define the quantity

$$h^*_{est}(\alpha, K) = \lim_{\epsilon \to 0} \limsup_{T \to \infty} \frac{1}{T} \log s^*_{est}(T, \epsilon, \alpha, K)$$

It was proved in [6] that $h_{est}(\alpha, K) = h^*_{est}(\alpha, K)$. In this paper, we deal with the following equation

$$\dot{x}(t) = A_{\sigma(t)} x(t), \quad x(0) \in K, \tag{2}$$

where $\sigma : \mathbb{R}_{\geq 0} \to \Sigma$, where $\#\Sigma < \infty$, is the *switching function* [3]. One should note that, after periodic sampling[1], equation (2) becomes a discrete LTV system. Therefore, equation (2) can be rewritten as

$$x_{k+1} = A_k x_k, \quad x(0) \in K, \tag{3}$$

where $A_k = \Phi(kT, 0)$, with $T > 0$ being the sampling period, and $\Phi(t, 0)$ being the *fundamental matrix* [2] of system (2). It is important to notice that $A_k \in \mathrm{Gl}(d, \mathbb{R})$, $\forall k \in \mathbb{N}$. In what follows, we will forget about the dynamics and work directly with the sequence of invertible matrices $(A_k)_{k \in \mathbb{N}}$.

Some remarks are in order. Since we deal only with LTV systems, we can drop the entropy's dependency on the set of initial conditions and write $h_{est}(\alpha, K) = h_{est}(\alpha)$, as long as K is compact with a nonempty interior, as pointed out in [13].

The following definitions were adapted from [5, 1, 11] and are reproduced here for the reader's convenience.

Definition 2.3. *A Lyapunov index of a sequence of matrices $(A_n)_{n \in \mathbb{N}}$ is a function $\lambda : \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$ with the following properties:*

- *$\lambda(\alpha v) = \lambda(v)$, for every $\alpha \neq 0$*

[1] This is also true if the sampling is aperiodic, however, we will keep it periodic in here for simplicity

- $\lambda(v+w) = \max\{\lambda(v), \lambda(w)\}$
- $\lambda(0) = -\infty$

Definition 2.4. *A Lyapunov exponent for a sequence of matrices $(A_n)_{n\in\mathbb{N}}$ is the following Lyapunov index*

$$\lambda(v) = \limsup_{n\to\infty} \frac{1}{n} \log\left(\left|\left|A^{(n)}v\right|\right|\right)$$

Definition 2.4 of Lyapunov exponent for sequences of matrices is compatible with the usual definition of Lyapunov exponents for linear dynamic systems [5], in the sense that the asymptotic exponential growth rate of the solution 3, i.e. $\limsup_{n\to\infty} \frac{1}{n}\log(||x_n||)$, with initial condition x_0 is given by $\lambda(x_0)$. Therefore, the Lyapunov exponents give an upper bound on the growth rate of the solutions of (3), and they will be used to obtain an expression for the entropy.

Also, note that, as a consequence of the two first itens in definition 2.3, a Lyapunov exponent of a sequence of matrices in $\mathrm{Gl}(d,\mathbb{R})$ can attain at most d distinct real values[2]. We convention that the $q \le d$ distinct real values attained by the Lyapunov exponents will be denoted by $\chi_i, \quad i = 1, \cdots q$, with the ordering $\chi_1 < \cdots < \chi_q$.

In order to give a justification for the previous definitions we present the following example.

Example 1. *Let $B_1 = \begin{bmatrix} \rho & 0 \\ 0 & \rho^{-1} \end{bmatrix}$ and $B_2 = \begin{bmatrix} \rho^{-1} & 0 \\ 0 & \rho \end{bmatrix}$. Consider the sequence $A_n = B_1$ whenever n is divisible by 4, and $A_n = B_2$ otherwise. Also let $\{e_1, e_2\}$ be the canonical basis for $\mathbb{R}^2$. Then it follows trivially that $\lambda(e_1) = -\frac{1}{2}\log\rho$ and $\lambda(e_2) = \frac{1}{2}\log\rho$*

A further consequence of definition 2.3 is that the Lyapunov exponents are constant over one dimensional subspaces except at the origin. Definition 2.3 gives us more as can be seen in the following definition.

Definition 2.5. *A filtration on $\mathbb{R}^d$ is a family of vector subspaces $\mathcal{V} = (E_i)_{i=0}^q$, with $q \le d$, such that $\{0\} = E_0 \subsetneq E_1 \subsetneq \cdots \subsetneq E_q = \mathbb{R}^d$.*

Further, we call $\mathbb{V} = \{v_i\}_{i=1}^d$ a normal basis of the filtration $\mathcal{V}$ if it is a basis for $\mathbb{R}^d$, and for every $j \ge 1$, the subset of $\mathbb{V}$ given by $\{v_i\}_{i=1}^{dim(E_j)}$ is a basis for E_j.

A filtration $\mathcal{V}_\lambda$ associated with the sequence of invertible matrices $(A_n)_{n\in\mathbb{N}}$ such that $E_i = \left\{v \in \mathbb{R}^d : \chi_i \ge \lambda(v)\right\}$, where λ is a Lyapunov index for the sequence, and χ_i are the Lyapunov exponente values of the sequence previously defined, is called an Oseledets filtration. Also, the subspaces $E_i \in \mathcal{V}_\lambda$ are called Oseledets subspaces. In addition, the dimension of E_i is called the multiplicity of the Lyapunov exponent value χ_i. Finally, define $\Lambda = \{\lambda_j\}_{j=1}^d$ as an ordered list with repetition where for every $j = 1, \cdots, d$, there exists some $i \in \{1, \cdots, q\}$ such that $\lambda_j = \chi_i$, and for every $i = 1, \cdots, q$, χ_i appears $dim(E_i)$ times in Λ. The order in Λ is arbitrary among those for which $\lambda_1 \le \cdots \le \lambda_d$. We call the elements $\lambda_i \in \Lambda$ the Lyapunov exponents with multiplicity of $(A_n)_{n\in\mathbb{N}}$.

[2]Note that, by convention [5], if we pick $v = 0$, $\lambda(v) = -\infty$

Another instrumental definition is that of *tempered sequence*

Definition 2.6. *A sequence $(A_n)_{n\in\mathbb{N}}$ is called tempered if*

$$\lim_{n\to\infty}\frac{1}{n}\log||A_n|| = 0$$

The idea of this definition is that, in this way, the growth rate of the sequence $(||A_n||)_{n\in\mathbb{N}}$ is, in some sense, slow. Note, however, that the Lyapunov exponents might be infinite. Take for instance $A_n = n$. Clearly, $\lim_{n\to\infty}\frac{1}{n}\log(A_n) = \lim_{n\to\infty}\frac{1}{n}\log(n) = 0$, but $A^{(n)} = n!$. Consequently, the Lyapunov exponent of this scalar sequence is $\lim_{n\to\infty}\frac{1}{n}\log(n!) = \infty$.[3]. Nonetheless, if $(||A_n||)_{n\in\mathbb{N}}$ is bounded, then the sequence of matrices is automatically tempered and its Lyapunov exponent is finite [14].

The central definition necessary for our discussion, however, is that of (Lyapunov) reularity. It imposes a direct relation between the growth rate of the volume of $\xi(K,t)$ as t increases, and the Lyapunov exponents.

Definition 2.7. *A sequence $(A_n)_{n\in\mathbb{N}}$ is called regular if*

$$\lim_{n\to\infty}\frac{1}{n}\log\left(\left|\det\left(A^{(n)}\right)\right|\right) = \sum_{i=1}^{d}\lambda_i$$

We call a discrete LTV system regular, if the sequence $(A_n)_{n\in\mathbb{N}}$ associated with it is regular. To illustrate the gist of this definition, we will resort to the following example of an irregular sequence.

Example 2. *Let B_1 and B_2 be as in Example 1. Consider the sequence $A_n = B_1$ if $n \in \{2^i, \cdots, 2^{i+1}\}$, for i odd, and $A_n = B_2$ otherwise. Note that $\det\left(A^{(n)}\right) = 1$, for all possible sequences $(A_n)_{n\in\mathbb{N}}$. Denote by $\{e_1, e_2\}$ the canonical basis. It follows that the $\lambda(e_1) > 0$ and $\lambda(e_2) > 0$, therefore the sequence cannot be regular.*

One should notice that in the previous example the limit superiors in the definition of Lyapunov exponents are not limits, i.e., they are different from the limit inferiors. That is actually the source of irregularity. As we will see in theorem 2.1, if the sequence is regular then those limits must exist.

The following theorem was extracted from [5] and synthesises equivalent characterizations for regularity for the class of tempered sequences.

Theorem 2.1. *Given a sequence $(A_n)_{n\in\mathbb{N}}$ of invertible and tempered matrices, let $\{v_1, \cdots, v_d\}$ be any normal basis associated with the Lyapunov exponents of the sequence, and let $\mathcal{I} \subset \{1, \cdots, d\}$ be any set of indices. Then, the following conditions are equivalent*

- $\lim_{n\to\infty}\frac{1}{n}\log\left(\det\left(A^{(n)}\right)\right) = \sum_{i=1}^{d}\lambda_i$;
- $\lim_{n\to\infty}\frac{1}{n}\log\left(\mathrm{vol}\left(\mathcal{P}_{A^{(n)}}\left(v_i : i \in \mathcal{I}\right)\right)\right) = \sum_{i\in\mathcal{I}}\lambda_i$.

[3]This part had the wrong statement, that the growth rate should be subexponential, in the published version of this work.

It is important to notice the special case of the second item of theorem 2.1 when $\mathcal{I}$ is a singleton, then we get that the limit superior in definition 2.4 can be replaced by the limit. Also notice that the example 1 is actually regular by the first item.

We will also need the following theorem 3.2.1 from [5]

Theorem 2.2. *Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of matrices in $Gl(d,\mathbb{R})$. Given an ordered normal basis $\{v_1,\cdots,v_d\}$ with respect to the Lyapunov exponent λ, there exists a sequence of orthogonal matrices $(U_n)_{n\in\mathbb{N}}$ such that:*

- *$C_m = U^*_{m+1}A_mU_m$ is upper triangular for each $m\in\mathbb{N}$;*
- *the columns of U_1 are the vectors $v_1,\cdots,v_d$;*
- *the sequence $(U_n)_{n\in\mathbb{N}}$ can be chosen such that the canonical basis $e_1,\cdots,e_n$ is normal with respect to the Lyapunov exponents $\tilde{\lambda}$ of $(C_n)_{n\in\mathbb{N}}$ and $\tilde{\lambda}(e_1)\leq\cdots\leq\tilde{\lambda}(e_n)$*

Note that orthogonal changes of coordinate do not change the values of the Lyapunov exponents, neither the regularity of a sequence, i.e., if $(A_n)_{n\in\mathbb{N}}$ is regular so is $(C_n)_{n\in\mathbb{N}}$ [5]. Also, this upper triangular reduction does not affect temperedness.

3 Estimation Entropy and Lyapunov Exponents

In this section, we consider the problem of finding a closed expression for the estimation entropy of discrete regular LTV systems. The following lemma is instrumental for what follows in this section.

Lemma 3.1. *Consider a regular matrix sequence $(A_n)_{n\in\mathbb{N}}\subset Gl(d,\mathbb{R})$, then the sequence $(B_n)_{n\in\mathbb{N}}$, defined by $B_n = A_ne^{-\alpha}$ is also regular. Moreover, any normal basis $\{v_1,\cdots,v_d\}$ for the Oseledets filtration of $(A_n)_{n\in\mathbb{N}}$ is also a normal basis for the Oseledets filtration of $(B_n)_{n\in\mathbb{N}}$, and the Lyapunov exponents λ_A associated with the sequence $(A_n)_{n\in\mathbb{N}}$, are such that $\lambda_B(v)=\lambda_A(v)+\alpha$, where λ_B are the Lyapunov exponents for the sequence $(B_n)_{n\in\mathbb{N}}$.*

Its proof is evident and will be omitted. Next, we prove that the estimation entropy of regular LTV systems can be written as a function of the Lyapunov exponents with multiplicity.

Theorem 3.1. *The estimation entropy of a regular discrete LTV system* (3) *is equal to $\sum_{i=1}^{d}\max\{0,\lambda_i+\alpha\}$. If the system is not regular, the estimation entropy is upper bounded by this quantity.*

Proof. We start by proving the lower bound. Consider $(A_n)_{n\in\mathbb{N}}$, a regular sequence of matrices that came from the system (3). Define $B_n = A_ne^{\alpha}$, it is true that $(B_n)_{n\in\mathbb{N}}$ is regular as well by Lemma 3.1. Also, by Lemma 3.1, any normal basis $\{v_1,\cdots,v_d\}$ with respect to the Lyapunov exponents λ_A of $(A_n)_{n\in\mathbb{N}}$ is also a normal basis with respect to the Lyapunov exponents λ_B of $(B_n)_{n\in\mathbb{N}}$. Moreover, $\lambda_B(v) =$

$\lambda_A(v)+\alpha$, for every $v \in \mathbb{R}^d$. As a remark, we denote by λ_i the distinct values of λ_A with multiplicity.

Hence, by the third item in Theorem 2.1, we know that for any ordered normal basis $\{v_1, \cdots, v_d\}$, and any set of indices $\mathcal{I} \subset \{1, \cdots, d\}$, we have that $\lim_{n\to\infty} \frac{1}{n} \log \operatorname{vol}\left(\mathcal{P}_{B^{(n)}}\{v_i : i \in \mathcal{I}\}\right) = \sum_{i\in\mathcal{I}} \lambda_B(v_i)$.

This implies that $\forall \delta > 0$, $\exists N \in \mathbb{N}$, such that $\forall n \geq N$

$$\operatorname{vol}\left(\mathcal{P}_{B^{(n)}}\{v_i : i \in \mathcal{I}\}\right) \geq e^{\sum_{i\in\mathcal{I}}(\lambda_B(v_i)-\delta)n} \tag{4}$$

Let $\mathcal{I}$ be the set of indices such that $\lambda_B(v_i) > 0$. Also, let $C = \{x_1, \cdots, x_N\}$ be an $(n, \epsilon, \alpha, U_\alpha)$- spanning set, where $U_\alpha = \operatorname{span}\{v_i, i \in \mathcal{I}\} \cap K$. Denote by $k_u = \#\mathcal{I}$.

Notice that, $\exists c_1 > 0$ and $\exists c_2 > 0$ such that $\mathcal{P}_{c_2 I_d}\{v_i : i \in \mathcal{I}\} \subset U_\alpha \subset \mathcal{P}_{c_1 I_d}\{v_i : i \in \mathcal{I}\}$, i.e., there are parallelepipeds, of the same dimension, that contain and are contained by the set U_α. Now, define $c(K,n) = \frac{\operatorname{vol}\left(B^{(n)}(U_\alpha)\right)}{\operatorname{vol}\left(\mathcal{P}_{B^{(n)}}\{v_i : i\in\mathcal{I}\}\right)}$. Notice that

$$\lim_{n\to\infty} \frac{1}{n} \log c(K,n) = 0,$$

which follows trivially from the inclusions above.

We also know that, $s^*_{est}(n, \epsilon, \alpha, U_\alpha) \geq \frac{\operatorname{vol}(B^{(n)}(U_\alpha))}{\operatorname{vol}(\mathbb{B}(B^{(n)}(x_i,n),\epsilon))} = \frac{\operatorname{vol}(B^{(n)}(U_\alpha))}{(2\epsilon)^{k_u}}$.

It is straight forward to see that $s^*_{est}(n, \epsilon, \alpha, K) \geq s^*_{est}(n, \epsilon, \alpha, U_\alpha)$, by the fact that any (n, ϵ, α, K)-spanning set is also a $(n, \epsilon, \alpha, U_\alpha)$-spanning set. By (4), we arrive at $s^*_{est}(n, \epsilon, \alpha, U_\alpha) \geq \frac{c(K,n)}{(2\epsilon)^{k_u}} e^{\sum_{i\in\mathcal{I}}(\lambda_B(v_i)-\delta)n}$.

Taking the logarithm, dividind by n, and taking lim sup we get that $h_{est}(\alpha)$ is lower bounded by $\limsup_{n\to\infty} \frac{1}{n} \log\left(c(K,n)\right) + \sum_{i\in\mathcal{I}}\left(\lambda_B(v_i) - \delta\right) - \frac{1}{n}\log(2\epsilon)^{k_u}$, which is equal to $\sum_{i\in\mathcal{I}}\left(\lambda_B(v_i) - \delta\right)$. Since δ was arbitrary, and we picked an arbitrary normal basis $h_{est}(\alpha) \geq \sum_{i\in\mathcal{I}} \lambda_B(v_i) = \sum_{i=1}^{d} \max\{0, \lambda_i + \alpha\}$

Now, we proceed to prove the upper bound on the estimation entropy. Note that we do not use the regularity property for this part of the proof.

Let $\{v_i : i = 1, \cdots, d\}$ be an ordered normal basis of the Oseledets' filtration of $(A_n)_{n\in\mathbb{N}}$. Note that, the basis $\{v_i : i = 1, \cdots, d\}$ can be chosen to be orthonormal by changing coordinates without loss of generality, so we will consider this for the rest of the proof. Take the parallelepiped $U = \left\{\sum_{i=1}^{d} \alpha_i v_i : \quad \alpha_i \in \left[0, \frac{\operatorname{diam}(K)}{||v_i||}\right]\right\}$ and note that $K \subset U$. Let I_i be the orthogonal projection of U on $\operatorname{span}\{v_i\}$, i.e., the side of the parallelepiped.

For fixed $\epsilon > 0$, fixed $\delta > 0$ and fixed $m \in \mathbb{N}$, divide each I_i in the following way: if $\lambda(v_i) \geq -\alpha$, divide it into subintevals of length $\frac{\epsilon e^{-(\lambda(v_i)+\alpha+\delta)m}}{d}$, otherwise divide it into subintervals of length $\frac{\epsilon}{d}$. Denote $k = \min\{j \in \mathbb{N} : \lambda_j \geq -\alpha\}$.

Now, define $\beta_i(x) = \frac{\langle x, v_i\rangle}{||v_i||}$, and $a_i = \min_{x\in U} \beta_i(x)$. We can now build the following grid $j_i = \left\lceil \frac{\beta_i(x)-a_i}{c_i} \right\rceil$, where $c_i = \frac{\epsilon e^{-(\lambda(v_i)+\alpha+\delta)m}}{d}$ for $i = k, \cdots, d$, and $c_i = \frac{\epsilon}{d}$ for $i = 1, \cdots, k-1$. In addition, let $j = (j_1, \cdots, j_d)$ be an index function. Also, define x_{j_i} as the $x \in I_i$ such that $\frac{\beta_i(x)-a_i}{c_i} = \frac{j_i}{2}$. Finally, $x_{\tilde{j}} = \sum x_{\tilde{j}_i}$, where $\tilde{j} = (\tilde{j}_i, \cdots, \tilde{j}_d)$.

Note that by construction, $\exists j \in \mathbb{N}^d$ such that $||x - x_j|| \leq \frac{\epsilon}{d}$ for $i \in \{1, \cdots, k-1\}$, and $||x - x_j|| \leq \frac{\epsilon e^{-(\lambda(v_i)+\alpha+\delta)n}}{d}$ for $i \in \{k, \cdots, d\}$. Finally, denote $C = \{x_1, \cdots, x_J\}$ the set of all possible x_j foredefined.

From the definition of limit superior, $\forall \delta > 0$, $\exists N \in \mathbb{N}$ such that $\forall n \geq N$ we have that $\left|\sup_{m \geq n}\left\{\frac{1}{m}\log\left(\left|A^{(m)}v_i\right|\right)\right\} - \lambda(v_i)\right| \leq \delta$. Then, we get $\frac{1}{n}\log\left(\left|A^{(n)}v_i\right|\right) - \lambda(v_i) \leq \sup_{m \geq n}\left\{\frac{1}{m}\log\left(\left|A^{(m)}v_i\right|\right)\right\} - \lambda(v_i) \leq \delta$. Finally, it implies that $\left|A^{(n)}v_i\right| \leq e^{(\lambda(v_i)+\delta)n}$, for all $n \geq N$.

Therefore, for $m \geq n \geq N$, and for $\delta \in (0, \lambda(v_{k-1}) - \alpha)$, it follows that:

$$
\begin{aligned}
\left|A^{(n)}x - A^{(n)}x_j\right| &= \left|\sum_{i=1}^{d} A^{(n)}\gamma_{i,j}v_i\right| \\
&\leq \sum_{i=k}^{d} e^{(\lambda(v_i)+\delta)n}\frac{\epsilon e^{-(\lambda(v_i)+\alpha+\delta)m}}{d} + \sum_{i=1}^{k-1}\frac{\epsilon}{d}e^{-(\alpha)n} = \epsilon e^{-\alpha n}
\end{aligned}
$$

for some j, where $\gamma_{i,j}$ is the orthogonal projection of $x - x_j$ over v_i.

Now, note that the collection $\mathcal{G}$ of functions $g_j(n) = A^{(n)}x_j$ satisfy the property that $\left|A^{(n)}x - g_j(n)\right| < \epsilon e^{-\alpha n}$ for $n \in \{N+1, \cdots, m\}$. Consider a family $\mathcal{F}(N)$ of functions $f_k : \{0, \cdots, N\} \to \mathbb{R}^d$, such that $\left|\left|A^{(t)}x - f_k(t)\right|\right| < \epsilon e^{-\alpha t}$, for $t \in \{0, \cdots, N\}$. Notice that, there always exist such a finite family for a finite N. Construct the functions $\hat{x}_p(t) = g_i(t)$ for $t \in \{0, \cdots, N\}$ and $\hat{x}_p(t) = f_k(t)$ for $t \in \{N+1, \cdots, m\}$, for every pair (i, k) possible. Define an index function $p : \mathbb{N}^2 \to \mathbb{N}$, that maps $(i, k) \mapsto p$. Hence, the set of functions $\hat{X} = \{\hat{x}_p\}$ constructed before is a (m, ϵ, α, K)-approximating set. It follows trivially that its cardinality is the product of the cardinality of $\mathcal{F}(N)$ and the cardinality of $\mathcal{G}$. For simplicity we will denote $\mathcal{F} = \mathcal{F}(N)$.

$$
\begin{aligned}
&\lim_{m\to\infty}\frac{1}{m}\log s_{est}(m, \epsilon, \alpha, K) \\
&\leq \lim_{m\to\infty}\frac{1}{m}\log\left(\#\mathcal{F}\prod_{i=k}^{d}\left(\frac{|I_i|de^{(\lambda_i+\alpha+\delta)m}}{\epsilon}\right)\prod_{i=1}^{k-1}\frac{|I_i|d}{\epsilon}\right) \\
&\leq \lim_{m\to\infty}\frac{1}{m}\log\left(\#\mathcal{F}\prod_{i=k}^{d}(|I_i|)\left(\frac{d}{\epsilon}\right)^d\right) + \sum_{i=k}^{d}(\lambda_i + \alpha + \delta) \\
&= \sum_{i=k}^{d}(\lambda_i + \alpha + \delta)
\end{aligned}
$$

For every fixed δ, $\mathcal{F}(N)$ is kept fixed, therefore the first term in the last inequality goes to 0. Also, from the fact that $\delta > 0$ can be chosen arbitrarily close to 0, it follows that $h_{est}(\alpha) \leq \sum_{i=1}^{d}\max\{0, \lambda_i + \alpha\}$ □

Therefore, we obtain a formula $h_{est}(\alpha) = \sum_{i=1}^{d}\max\{0, \lambda_i + \alpha\}$ for the estimation entropy. It should be remarked that this expression is closely related to Pesin's formula, since they are essentially the same when $\alpha = 0$. However, the conditions under which Pesin's formula holds are very different then those required here [7]. Also, there is no direct use of measure theoretical concepts in the present derivation.

4 Sufficient conditions for regularity

In this section we advocate in favor of regular systems, providing several classes of systems that are inherently regular. We start claiming that continuous-time regular systems give rise to discrete-time regular systems under sampling, then present the case of systems with periodic switching. Next, we present the very interesting case o simultaneously triangularizable systems, and give sufficient conditions on the average activation time for each one of the modes. We conclude with a class of systems that preserve a given probability measure such as Markov Jump Linear Systems (MJLS).

4.1 Sampled Regular Systems

Consider the continuous-time linear system described by equation

$$\dot{x}(t) = A(t)x(t). \tag{5}$$

One can define regularity for system (5) as in the discrete case, using the following definition of Lyapunov exponents $\lambda(v) = \limsup_{t\to\infty} \frac{1}{t} \log ||\Phi_A(t,0)v||$, where $\Phi_A(t,0)$ is the fundamental solution of system (5). Then system (5) is regular if $\lim_{t\to\infty} \frac{1}{t} \log |\det(\Phi_A(t,0))| = \sum_{i=1}^{d} \lambda_i$.

The following proposition shows that the discrete-time system that originates from sampling a regular continuous-time system is regular.

Proposition 4.1. *Given a continuous-time LTV system as in equation* (5)*. Define* $x_n = x(T_p n)$ *and* $A_n = \Phi(nT_p, (n-1)T_p)$ *for* $n \in \mathbb{N}$*, where* $\Phi(t,0)$ *is the fundamental matrix of* (5), and T_p is the sampling time. Then, sequence $(A_n)_{n\in\mathbb{N}}$ is regular.

Proof. Denote by $\bar{\lambda}$ the Lyapunov exponent of the sampled sequence $(A_n)_{n\in\mathbb{N}}$, and by λ the Lyapunov exponent of the continuous-time system. Then,

$$\lambda(v) = \limsup_{t\to\infty} \frac{1}{t} \log ||\Phi_A(t,0)v|| \geq \limsup_{n\to\infty} \frac{1}{nT_p} \log ||\Phi(nT_p,0)v||,$$

also, from the fact that (2) is regular, it follows that the lim sup is equal to the lim inf

$$\lambda(v) = \liminf_{t\to\infty} \frac{1}{t} \log ||\Phi_A(t,0)v|| \leq \liminf_{t\to\infty} \frac{1}{nT_p} \log ||\Phi(nT_p,0)v||,$$

hence, $\bar{\lambda}(v) = \lim_{n\to\infty} \frac{1}{n} \log || \left(A^{(n)}v\right) || = T_p \lambda(v)..$ □

One more consideration must be made regarding temperedness. We define temperedness for continuous-time systems in an analogous way as for discrete-time systems as $\lim_{t\to\infty} \frac{1}{t} \log ||A(t)|| = 0$. From this definition, it is immediate that a sampled linear tempered system is also tempered.

4.2 Periodic case

In this subsection we will deal with continuous-time systems and obtain the result for discrete-time systems as a consequence by using the previous result on sampled systems. Given a LTV system $\dot{x} = A(t)x$, such that $A(t) = A(t+T)$ for some $T > 0$ and for all $t \in \mathbb{R}$, one can write the fundamental matrix of it as [2] $\Phi(t,0) = R(t)e^{tF}$, where, $R(t)$ is a periodic matrix with period T, and e^{TF} is the monodromy matrix. In this case the real part of the eigenvalues of F are the Lyapunov exponents of the system, also known as Floquet exponents. Now, noticing that $\Phi(0,0) = I_d$ one gets that $R(T) = I_d$. Therefore $\Phi(kT,0) = e^{kTF}$.

One can now take any, possibly complex valued, eigenvector v_i associated with the i-th eigenvalue $\tilde{\lambda}_i$, with real part λ_i, of F and get

$$\lambda(v_i) = \limsup \frac{1}{k} \log\left(||\Phi(kT,0)v_i||\right) = \lambda_i,$$

which shows that the real part of the eigenvalues of F are the Lyapunov exponents.

From the fact that $R(t)$ is periodic, one concludes that $\Phi(t,0)$ is bounded, which implies that it is a tempered system. Moreover, from Proposition 4.1 we conclude that any sampling of the fundamental matrix, even if the sampled system is not periodic itself, renders the sampled system regular. The discrete-time case is a subcase of the continuous-time one.

4.3 Upper triangularizable matrices

The case of simultaneously triangularizable matrices $(A_n)_{n\in\mathbb{N}}$ has been studied in [13], where upper and lower bounds for the topological entropy were given. We show here that, for a finite set of modes, under the hypothesis of regularity the bounds given in [13] are not tight in general. Moreover, we present a closed form expression for the estimation entropy of these systems. In what follows, we consider that the sequence of matrices is already in triangular form, since a linear change of coordinates does not affect the result.

Theorem 4.1. *Consider a sequence of triangular matrices* $(A_n)_{n\in\mathbb{N}}$*, such that* $A_n = B_j$*, for some* $B_j \in \mathcal{B}$*,* $j = 1, \cdots, m$*, where* $\mathcal{B} = \{B_1, \cdots, B_m\} \subset Gl(\mathbb{R}, d)$ *is a set of upper triangular matrices of cardinality* m*. Also cosider* $\tau_i(n) = \sum_{j=0}^{n} \mathbb{I}_{A_j = B_i}$ *be the activation time. Then, a sufficient condition for regularity is that the average activation time* $\frac{\tau_i(n)}{n}$ *of each mode* i *converges to some value* $\lim_{n\to\infty} \frac{\tau_i(n)}{n} = \rho_i$*.*

In that case, the estimation entropy is given by

$$h_{est}(\alpha) = \sum_{i=1}^{d} \max\left\{0, \sum_{j=1}^{m} \rho_j \log{(B_j)}_{ii} + \alpha\right\} \tag{6}$$

Proof. A sufficient condition for regularity of a sequence of upper triangular matrices proved in Theorem 3.1.3 from [5] is that the product of the logarithm of the diagonal entries converge. These diagonal entries are

$$\left(A^{(n)}\right)_{ii} = \prod_{j=1}^{m} (B_j)_{ii}^{\tau_j(n)}\, e^{\alpha n},$$

Now, taking the logarithm and dividing by n we obtain that

$$\begin{aligned}\frac{1}{n}\log\left(A^{(n)}\right)_{ii} &= \frac{1}{n}\log\left(\prod_{j=1}^{m}(B_j)_{ii}^{\tau_j(n)}\right)+\alpha\\ &= \sum_{j=1}^{m}\frac{\tau_j(n)}{n}\log(B_j)_{ii}+\alpha.\end{aligned}$$

It follows then that if $\frac{\tau_j(n)}{n} \to \rho_j$, then $\lim_{n\to\infty}\frac{1}{n}\log\left(A^{(n)}\right)_{ii}$ exists and equals $\sum_{j=1}^{m}\rho_j\log(B_j)_{ii}+\alpha$.

We know that, for an upper triangular sequence of matrices, whenever $\lim_{n\to\infty}\frac{1}{n}\left(A^{(n)}\right)_{ii}$ converges it converges to the values of the Lyapunov exponents with multiplicities by the aforementioned Theorem 3.1.3 of [5]. Therefore, by the formula $h_{est}(\alpha) = \sum_{i=1}^{d}\max\{0, \lambda_i+\alpha\}$, we have the desired result.

□

It is important to remark that if B_j is the exponential of an upper triangular matrix C_j, then $\log(B_j)_{ii} = (C_j)$. It follows that by using the fact that the sum of the maximum is less than or equal to the maximum of the sums and taking $\alpha = 0$, we recover on the left hand side the lower bound for the topological entropy of linear switched systems $\max\left\{0, \sum_{j=1}^{m}\rho_j\left(\mathrm{Tr}(C_j)\right)\right\}$ presented in theorem 4 of [13]. Therefore, under the mild assumption of regularity, we improve the result previously reported in the literature.

4.4 Oseledets theorem

One may ask how common of a property regularity is. In this section we show that regularity is a generic property. The sense in which regularity is general is a measure theoretic one to be made precise later. Before discussing this result we need some definitions.

Definition 4.1 (Linear Cocycle [11])**.** *Let $(M, \mathscr{B}, \mu)$ be a probability space, $f : M \to M$ be a measure-preserving map. Let $A : M \to Gl(\mathbb{R}, d)$. The linear cocycle defined by A over f is the transformation $F : M\times\mathbb{R}^d \to M\times\mathbb{R}^d$ with $F(x, v) = (f(x), A(x)v)$. It follows that $F^n(x, v) = (f^n(x), A(x)v)$ for every $n \geq 1$. Moreover, if f is invertible, then so is F, with inverse $F^{-1}(x, v) = (f^{-1}(x), A^{-1}(x))$.*

To clarify the discussion it is useful to bear in mind the two following examples. In what follows, let $Y = \{B_1, \cdots, B_m\}$ be the ordered set of modes, let $\mathcal{Y} = 2^Y$ be a σ-algebra. Further, let $f : M \to M$ be the shift map, i.e., $(\alpha_k)_{k\in\mathbb{N}} \mapsto (\alpha_{k+1})_{k\in\mathbb{N}}$ and consider $A : M \to \mathrm{Gl}(\mathbb{R}, d)$, with $(\alpha_k)_{k\in\mathbb{N}} \mapsto \alpha_0$. Finally, let $F : M\times\mathbb{R}^d \to M\times\mathbb{R}^d$ be the linear cocycle defined by A over f. Note that $F^n((\alpha_k)_{k\in\mathbb{N}}, v) = ((\alpha_{k+n})_{k\in\mathbb{N}}, \alpha_{n-1}\cdots\alpha_0 v)$. We will define M, $\mathscr{B}$ and μ in the examples.

Example 3 (Bernoulli Shifts [11]). *For this example, let $M = Y^{\mathbb{N}}$ and $\mathscr{B}$ be the product σ-algebra. Next, we need to introduce a measure on M. To do that, we put a probability measure $p : \mathcal{Y} \to [0,1]$ on Y and consider the product measure defined by $\mu(\{(\alpha_k)_{k\in\mathbb{N}} : \alpha_i \in E_i, \cdots, \alpha_j \in E_j\}) = p(E_i)\cdots p(E_j)$ for every $i \leq j$ and any sets $E_i, \cdots, E_j \subset Y$. Note that μ is invariant under the shift map. Note that this is the probability measure induced by a sequence of i.i.d. trials of a Bernoulli process.*

Example 4 (Markov Shifts [12]). *Let $P = (p_{i,j}) \in \mathcal{M}(m, \mathbb{R})$ be the transition matrix of an irreducible and aperiodic discrete-time discrete-state Markov chain, that represents the switching of the modes $B_i \in Y$. Assume that $p(B_i)$, for $B_i \in Y$ is the unique stationary measure of this chain. Also, let $M \subset Y^{\mathbb{N}}$ be the set with all the sequences of matrices that are valid with respect to the Markov chain.*

Now, we need to introduce a measure on M. Consider the measure defined as $\mu(\{(\alpha_k)_{k\in\mathbb{N}} : \alpha_i \in E_i, \cdots, \alpha_j \in E_j\}) = p_{\pi(\alpha_{j-1}),\pi(\alpha_j)} \cdots p_{\pi(\alpha_i),\pi(\alpha_{i+1})} p(\alpha_i)$, for every $i \leq j$ and any $\mathcal{Y}$-measurable sets $E_i \subset Y$, where $\pi : Y \to \#Y$ returns the index of the corresponding mode[4]. Note that μ is invariant under the shift map. Also, since P is irreducible, f is ergodic, see for instance theorem 1.13 from [12].

Consider a discrete-time linear switched system with distinct modes $B_i \in \mathcal{B}$, $i \in \{1, \cdots, m\}$. Then, the switching function $\sigma : \mathbb{N} \to 1, \cdots, m$ gives rise to a bijective correspondence between $\sigma(\mathbb{N}) \subset \{1, \cdots, m\}^{\mathbb{N}}$ and $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}^{\mathbb{N}}$ through the equality $A_n = B_{\sigma(n)}$. We call $\sigma(\mathbb{N})$ a switching and we denote by $\Sigma \subset \{1, \cdots, m\}^{\mathbb{N}}$ the set of valid switches, e.g. for the Markov example $\Sigma = \left\{(x_n)_{n\in\mathbb{N}} \in \{1, \cdots, m\}^{\mathbb{N}} : p_{\pi(x_{j+1}),\pi(x_j)} > 0, \forall j \in \mathbb{N}\right\}$, where x_k is the index of the active mode at time k.Also, because of the bijection, we say that the measure μ in M is a measure in Σ.

Now, we can present the Oseledets' theorem.

Theorem 4.2 (Oseledets [11, 5]). *Let $(M, \mathscr{B}, \mu)$ be a probability space, $f : M \to M$ be an invertible measure-preserving map. Let $A : M \to Gl(\mathbb{R}, d)$ be such that $\log^+ ||A|| \in L^1(\mu)$ and $\log^+ ||(A)^{-1}|| \in L^1(\mu)$. Also consider the linear cocycle defined by A over f.*

Then, for μ-almost every $x \in M$, there is $k = k(x)$, numbers $\lambda_k(x) > \cdots > \lambda_1(x)$ and a filtration $\{0\} = E_x^1 \subsetneq \cdots \subsetneq E_x^k = \mathbb{R}^d$ such that, $\forall i = 1, \cdots, k$:

- $k(f(x)) = k(x)$ *and* $\lambda_i(f(x)) = \lambda_i(x)$ *and* $A(x)(E_x^i) = E_{f(x)}^i$;
- $\lim_{n\to\infty} \frac{1}{n} \log ||A^{(n)}(x)v|| = \lambda_i(x)$, *for all* $v \in E_x^{i+1} \setminus E_x^i$, *with* $E_x^1 = \{0\}$,

if f is ergodic, the multiplicities of the Lyapunov exponents $k(x)$ is constant and, consequentely, the dimension of the subspaces E_x^i, also $\lambda_i(x) = \lambda_i$ is constant a.e.. We denote by C the measurable set with $\mu(C) = 1$ on which this theorem is true and call it the set of regular realizations.

[4]The index of an element of $B_i \in Y$ is i.

Oseledets' theorem ensures that for any shift invariant measure μ we have a full measure set C_μ in Σ, such that all its switches are regular. So the set of all regular switches contains the union $\cup_\mu C_\mu$ such that μ is shift invariant. In this sense, the set of all regular switchings is generic.

One important case is the aforementioned periodic switching case. Using Oseledets' theorem, we can rederive the result from Subsection 4.2 by taking $\mu = \frac{1}{T}\sum_{i=0}^{T-1} \delta_{f^i\left((A_n)_{n\in\mathbb{N}}\right)}$, where δ_x is the so-called Dirac measure at x [12]. It is clear that μ is shift invariant and ergodic. Moreover, the $\log^+(||A||^{\pm}) \in L^1(\mu)$ condition is satisfied if the sequence $(A_n)_{n\in\mathbb{N}}$ is bounded, so it trivially holds. Therefore, every periodic switching is regular. Another important case is that of Markov Jump Linear Systems, which are modeled, in our framework, as in Example 4. Therefore, ergodic MJLS have realizations that are, with probability one, regular.

5 Conclusion and Future works

In this paper we presented a closed expression for the estimation entropy for a class of linear switched systems, expanding and improving results already reported in the literature. We showed that the class of tempered regular systems contains several examples of practical interest such as periodic switched systems, simultaneuously triangularizable systems, and randomly switched systems, in particular Markov Jump Linear Systems.

As future works we propose to present a quantization method that achieves the estimation entropy for the class of tempered systems regardless of the fact that the system is regular or not. Also, it remains an open question if there is a general closed expression for irregular systems, future works should address that problem. The authors also belive that the study of almost periodic nonlinear oscillators can be better understood though linearization over the periodic trajectory and using the framework of linear time varying systems presented here. Finally, the study of control of linear switched systems with minimum average data-rate has yet to be properly considered.

6 Acknowledgements

This work was supported by the NSF grant CMMI-1662708 and the AFOSR grant FA9550-17-1-0236.

References

[1] L. Arnold, Random Dynamical Systems, *Springer Berlin Heidelberg*, 1998

[2] R.W. Brockett, Finite Dimensional Linear Systems, *Society for Industrial and Applied Mathematics*, (SIAM), Classics in Applied Mathematics, 2015

[3] D. Liberzon, Switching in Systems and Control, *Birkhäuser Boston*, Systems & Control: Foundations & Applications, 2003

[4] L. Barreira and Y. Pesin, Nonuniform Hyperbolicity: Dynamics of Systems with Nonzero Lyapunov Exponents, *Cambridge University Press*, 2014.

[5] L. Barreira, Lyapunov Exponents, *Springer International Publishing*, 2017.

[6] D. Liberzon and S. Mitra, Entropy and Minimal Bit Rates for State Estimation and Model Detection, IEEE Transactions on Automatic Control, vol 63, n 10, pp 3330-3344, 2018

[7] R. Mañé, A proof of Pesin's formula, *Cambridge University Press (CUP)*, Ergodic Theory and Dynamical Systems, vol. 1, pp 95-102, 1981.

[8] A.S. Matveev and A.V. Savkin, Estimation and Control over Communication Networks, *Birkhauser*, 2007.

[9] G.N. Nair and R.J. Evans and I.M.Y. Mareels and W. Moran, Topological Feedback Entropy and Nonlinear Stabilization, IEEE Transactions on Automatic Control, vol. 49, n 9, pp 1585-1597, 2004.

[10] H. Sibai and S. Mitra, Optimal Data Rate for State Estimation of Switched Nonlinear Systems, *ACM Press* , Proceedings of the 20th International Conference on Hybrid Systems: Computation and Control, pp 71-80, 2017.

[11] M. Viana, Lectures on Lyapunov Exponents, *Cambridge University Press*, Cambridge Studies in Advanced Mathematics, 2014.

[12] P. Walters, An Introduction to Ergodic Theory, *Springer New York*, 2000.

[13] G. Yang and A. J. Schmidt and D. Liberzon, On Topological Entropy of Switched Linear Systems with Diagonal, Triangular, and General Matrices, *IEEE*, IEEE Conference on Decision and Control (CDC), 2018

[14] L. Barreira and C. Valls, Lyapunov regularity via singular values, *Transactions of the American Mathematical Society*, vol 369, n 12, pp 8409 - 8436, 2017