Is switching systems stability harder for continuous time systems?

Vladimir Y. Protasov and Raphaël M. Jungers

Abstract—We analyse the problem of stability of a continuous-time linear switching system (LSS) versus the stability of its Euler discretization. It is well-known that the existence of a positive τ for which the corresponding discrete-time system with stepsize τ is stable implies the stability of the LSS. Our main goal is to obtain a converse statement, that is to estimate the discretization stepsize τ > 0 up to a given accuracy ε > 0. This would lead to a method for deciding the stability of a continuous time LSS with a guaranteed accuracy. As a first step towards the solution of this problem, we show that for systems of matrices with real spectrum the parameter τ can be effectively estimated. We prove that in this special case, the discretized system is stable if and only if the Lyapunov exponent of the LSS is smaller than −Cτ, where C is an effective constant depending on the system. The proofs are based on applying Markov-Bernstein type inequalities for systems of exponents.

I. INTRODUCTION

Switching linear systems have been at the center of great attention in the past years. In continuous time, these are systems that satisfy the following equation:

\[
\begin{align*}
\dot{x}(t) &= A(t)x(t); \\
x(0) &= x_0; \\
A(t) &\in \mathcal{A}, \quad t \geq 0,
\end{align*}
\]

where \( \mathcal{A} \) is here chosen to be a finite set of matrices. Let alone the theoretical challenges they offer, they appear in many applications of high importance nowadays, like viral disease treatments optimization [5], Multi-hop Control Networks [14], multi-agent and consensus systems [2].

More generally, they provide a neat and well-defined framework for developing theoretical and computational tools, that can be used as building blocks for analyzing or designing more complex systems, like more general hybrid systems, systems with quantized signals, or event-triggered control schemes. See [13], [20], [21], [30] for general surveys on techniques and applications of switching systems.

In this paper we focus on the stability analysis of switching systems.

Definition 1: The system (1) is stable if \( \|x(t)\| \to 0 \) as \( t \to +\infty \) for every \( x_0 \) and any measurable function \( A(t) \).

In the above definition (as in the rest of the paper, unless explicitly stated), the norm can be chosen arbitrarily. The problem of deciding whether a switching system is stable has been studied in many papers both in discrete-time (see, e.g. [4], [11], [15], [26], [29]) and continuous-time (e.g. [1], [7], [12], [22]). This stability is ruled by the so-called worst-case Lyapunov exponent (abbr. Lyapunov exponent), which we now define.

Definition 2: The (worst case) Lyapunov exponent \( \sigma(A) \) of a switching system defined by a set of matrices \( A \) is the infimum of the numbers \( \alpha \) such that \( \|x(t)\| \leq Ce^{\alpha t} \) for every trajectory of (1).

Obviously, if \( \sigma < 0 \), then the system is stable. The converse is also true: if the system is stable, then \( \sigma < 0 \) (see for instance [23], [24]). Thus, the stability of the LSS defined by the set of matrices \( A \) is equivalent to the condition \( \sigma(A) < 0 \). Similar definitions exist for discrete-time switching linear systems, where the quantity \( \rho(A) = \exp(\sigma(A)) \) is called the joint spectral radius of the set \( A \) (see Definition 4 below).

Surprisingly, it seems that the state of the art is quite more advanced in the discrete time framework.

It is known that the question of approximating the joint spectral radius up to a guaranteed accuracy is NP-hard [32], but algorithms are known that perform this approximation up to a desired accuracy, in a time bounded by an (exponential) function of the dimension and the inverse of the desired accuracy. More precisely, there exist algorithms with the following properties:

Definition 3: An algorithm for computing the Lyapunov exponent with guaranteed accuracy is an algorithm which, given any \( \epsilon > 0 \), and any set of \( m \) matrices \( A = \{A_1, \ldots, A_m\} \subset \mathbb{R}^{d \times d} \), returns an approximation of the Lyapunov exponent which is within a relative accuracy of \( \epsilon \) in less than \( f(m, \epsilon, d, ||A||) \) computer operations, where \( f \) is a computable function, and

\[
||A|| \triangleq \max_{A_i \in A} \{||A_i||\}.
\]

See for instance [13, Theorem 2.12] for such an algorithm. On the other hand, no such result is known for continuous-time switching systems. The stability analysis of continuous time LSS is also NP-hard [12], but in fact no method is known that approximates the Lyapunov exponent, even in exponential time, up to a guaranteed accuracy. The methods proposed in the literature are, to the best of our knowledge, only sufficient conditions for stability. (Often, they are stated as LMIs and can be checked with SemiDefinite Programming, see [7], [8].)

Such approximation algorithms are important for several reasons. Suppose that one makes use of a sufficient
stability condition, and that the set $A$ does not satisfy that condition. One cannot conclude anything about his system, so that the system could in fact be highly stable (i.e. $\sigma(A) << 0$) but yet, could violate the condition. On the contrary, an algorithm with guaranteed accuracy ensures that one is able to conclude stability provided that $\sigma$ is small enough. On the theoretical side, it is also interesting to have such a function $f$, because it provides a natural way to compare different methods with each other in a clear and unambiguous way. With methods that rely on conditions that are only sufficient conditions for stability on the other hand, one is bound to provide particular instances (i.e., sets of matrices) for which old conditions fail to prove stability, while newly provided ones actually do.

The problem of finding such an algorithm is hard. In this paper, we make the first step towards its solution and tackle the particular case where all the matrices defining the system have real spectra. In this case, we are able to compute the stepsize $\tau$ such that the discretized switching system (i.e., defined by $I + \tau A = \{(I + \tau A : A \in A)\}$ is guaranteed to have the joint spectral radius close to $\exp(\sigma(A))$, that is, such that one can rely on well-known methods for discrete-time switching systems stability analysis in order to decide stability of the continuous-time system. We believe that the same result holds for general systems as well, although our method of proof only works for matrices with real spectra.

The remainder of the paper is as follows: In Section II, we expose the main idea of our method and formulate two fundamental theorems. Their proofs are in Sections III and IV.

II. MAIN RESULTS AND IDEAS

A. The mathematical setting

We consider a linear switching system (LSS) of the form (1). If $A(t)$ is a summable function that takes values in a given finite set of $d \times d$-matrices $A$, then $x(t)$ is a univariate vector-function taking values in $\mathbb{R}^d$ and this function belongs to the Sobolev space $W_1^2([0,a], I$ is the identity matrix, $\text{sp}(A)$ is the spectrum of the matrix $A$, $\rho(A)$ is the spectral radius (the largest modulus of eigenvalues), $e = (1, \ldots, 1) \in \mathbb{R}^d$ is the vector of ones.

As said above, our strategy to approximate the Lyapunov exponent will be to pass to the corresponding discrete time system with the step $\tau > 0$:

$$X(k+1) = X(k) + \tau A(k) X(k), \quad A(k) \in A, \forall k > 0. \quad (2)$$

This system is obtained from (1) by denoting $X(k) \approx x(k\tau)$ and by replacing the derivative $\dot{x}(t)$ by the divided difference $\frac{x((k+1)\tau) - x(k\tau)}{\tau}$. The stability of the discrete system is equivalent to the inequality $\rho(I + \tau A) < 1$, where $\rho(I + \tau A)$ is the joint spectral radius of the family

$$I + \tau A = \{I + \tau A \mid A \in A\}.$$

**Definition 4 (Joint Spectral Radius – JSR [29]):** the joint spectral radius of a set of matrices $\mathcal{M}$ is defined as

$$\rho(\mathcal{M}) = \lim_{k \to \infty} \max_{A_1, \ldots, A_k \in \mathcal{M}} ||A_1 \ldots A_k||^{1/k}, \quad (3)$$

where $|| \cdot ||$ is any matrix norm on $\mathbb{R}^{n \times n}$.

There has been a large research effort in the last years in order to efficiently compute the Joint Spectral Radius [9], [13], [25]. As said above, there exist approximation algorithms with guaranteed accuracy for the joint spectral radius. (A recent line of work [10], [15] even presents algorithms that find the exact value (in the form of a root of some polynomial) in a vast majority of cases.) This makes it possible to determine the stability of discrete time LSS efficiently. Thus, our goal is to apply these efficient methods to improve the state of the art in continuous time.

How to compute a valid discretization step $\tau$? The following well-known fact is central in our method:

**Theorem 1:** [23], [24] If the discrete time system (2) is stable for some $\tau > 0$, then it is stable for all smaller $\tau$, and the corresponding continuous system (1) is stable.

In other words, the inequality $\rho(I + \tau A) < 1$ for some $\tau > 0$ implies the stability of the continuous time LSS (1). The converse is also true:

**Theorem 2:** [23], [24] If the continuous time LSS is stable, then there is $\tau > 0$ such that $\rho(I + \tau A) < 1$.

However, the practical implementation of this result may be hard, because it is a priori not clear how small $\tau$ should be. If we take, say, $\tau = 10^{-3}$ and get $\rho(I + \tau A) > 1$, then no conclusion can be drawn on the stability of the continuous LSS. Our goal is to estimate the parameter $\tau$ for a given family $A$ such that the quantity $\rho(I + \tau A)$ is close enough to $\exp(\sigma(A))$. If we knew such a bound for $\tau$, then we could determine the stability of a continuous time system by merely checking the stability of the discretized system with a single step length $\tau$. Of course, this problem is well-defined provided the stability is determined with a finite accuracy $\varepsilon > 0$. Indeed, when $\varepsilon$ tends closer to zero, the critical value for $\tau$ gets arbitrarily small. To summarize, what we are up to is a numerical function as the following:

**Definition 5:** For any two positive real numbers $a, \varepsilon > 0$, we define the function $s(a, \varepsilon)$ as the largest positive number such that, for every $\tau < s(a, \varepsilon)$, for any set of matrices $A$ such that $r(A) \triangleq \max \{\rho(A) : A \in A\} = a$, the inequality $\rho(I + \tau A) \geq 1$ implies that $\sigma(A) \geq -\varepsilon$.

This definition may seem puzzling, because there is no evidence that this number depends on $r(A)$. It turns out, however, that our bound actually depends on this quantity. The only thing that matters, yet, is that the quantity is effectively computable, which is definitely the case with $r(A)$.

Thus, for every $\tau < s(r(A), \varepsilon)$ we have: if $\rho(I + \tau A) < 1$, then $\sigma < 0$ (the LSS $\dot{x} = Ax$, $A(t) \in A$ is stable); if $\rho(I +
\[ \tau A \geq 1, \text{ then } \sigma \geq -\varepsilon \text{ (the LSS } \dot{x} = (A + \varepsilon I)x, A(t) \in A \text{ is unstable).} \]

A lower bound for \( s(r(A), \varepsilon) \) may be derived from the local Lipschitz continuity of the joint spectral radius. However, the proofs of existence of the Lipschitz constant for the joint spectral radius (\cite{18}, \cite{26}) do not allow to bound this constant efficiently, so that we were not able to lower bound the function \( s(a, \varepsilon) \) with this technique. This is, in a sense, unavoidable, because the Lipschitz constant may be very large already even for one matrix (this can happen if the matrix has two close eigenvectors). It turns out, however, that this situation does not occur for \( s(a, \varepsilon) \).

**B. Main theorems**

The following theorem allows to pick a discretization step providing an accuracy \( \varepsilon \), given the dimension \( d \) and the maximal spectral radius \( r \) of the matrices.

**Theorem 3:** If all matrices of \( A \) have real spectrum, then

\[ s(a, \varepsilon) \geq \frac{6 \varepsilon}{(16 d^2 - 24 d + 11) r^2}. \]

**Corollary 1:** Assume all matrices of \( A \) have real spectrum; then if if the discrete time system with the step length

\[ \tau = \frac{6 \varepsilon}{(16 d^2 - 24 d + 11) r^2} \tag{4} \]

is not stable, then \( \sigma(A) \geq -\varepsilon \), i.e., the continuous time system with the set of matrices \( A + \varepsilon I \) is not stable.

We see that the lower bound \( s(a, \varepsilon) \) for the critical value of \( \tau \) is linear in \( \varepsilon \), which is natural, and decays with the dimension as \( d^{-2} \), which is much better than one could expect.

Theorem 3 follows from our next result, Theorem 4. First we need to introduce some more notation. For every \( d \geq 2 \) there is a unique polynomial \( b_d(t) \) of degree \( d - 1 \) which solves the following extremal problem: among all algebraic polynomials \( p(t) \) of degree \( d - 1 \) such that \( \|e^{-t}p(t)\|_{\mathbb{C}(\mathbb{R}_+)} \leq 1 \), find the maximal value of \( |p'(0)| \) (here \( \|p(t)\|_{\mathbb{C}(\mathbb{R}_+)} \) denotes the maximal value of \( |p(t)| \) over \( \mathbb{R}_+ \)). The extremal polynomial \( b_d(t) \) is called the Tchebychev polynomial with the Laguerre weight. We call the function \( S_d(t) = e^{-t}b_d(t) \) the L-Tchebychev polynomial. This function has \( d \) points of alternance on \( \mathbb{R}_+ \), i.e., there are \( d \) numbers \( 0 = \nu_1 < \nu_2 < \cdots < \nu_d \) such that \( S'(\nu_k) = ( -1 )^k \) and \( S''(\nu_k) = 0 \). The corresponding algebraic polynomial \( b_d = b \) is characterized by the equalities \( b(0) = -1, b(\nu_k) = b'(\nu_k) = ( -1 )^k e^{\nu_k}, k = 2, \ldots, d \) \cite{3}.

**Theorem 4:** For every family \( A \) of matrices with real spectrum and for every \( \varepsilon > 0 \), we have

\[ s(a, \varepsilon) \geq \frac{2 \varepsilon}{r^2 |S''(0)|}. \]

The function \( S_d \) can be evaluated numerically, the values of its derivatives at zero were listed in \cite{31} for all \( d = 1, \ldots, 20 \). In table 1 we write \( |S''(0)| \) for \( d \leq 10 \).

| \( d \) | \( |S''(0)| \) |
|---|---|
| 2 | 8.18224 |
| 3 | 25.1574 |
| 4 | 52.5873 |
| 5 | 90.5857 |
| 6 | 139.191 |
| 7 | 198.420 |
| 8 | 268.283 |
| 9 | 348.788 |
| 10 | 439.938 |

**TABLE I**

The values of \( |S''(0)| \) for \( d = 2, \ldots, 10 \)

Substituting in Theorem 4 we get the lower bounds shown in table 2.

These numerical bounds are better than the general bound in Theorem 3 for these \( d \). The reason is that they are based on the actual value of \( |S''(0)| \) which can be numerically computed, while Theorem 3 only uses a general bound on this quantity (from the main result of \cite{31} on the Markov-Bernstein type inequality for the polynomials with the Laguerre weight), which is valid in arbitrary dimension. We show that in the next section. Then, in the following section we give a proof of Theorem 4.

**Example 1:** Let us consider the set of matrices

\[ A = \left\{ \begin{pmatrix} -0.1322 & 0.0349 & -0.1182 \\ 0.0953 & -0.1397 & -0.1719 \\ 0.0787 & 0.0223 & -0.3281 \end{pmatrix}, \begin{pmatrix} 0.0891 & 0.1397 & -0.0916 \\ 0.0338 & -0.2269 & -0.0707 \\ 0.7417 & 0.3028 & -0.5121 \end{pmatrix} \right\}. \tag{5} \]

Both matrices happen to have real spectra, and hence we can apply our results. We chose an accuracy \( \varepsilon = 0.2 \). Applying Corollary 1 and the estimate from Table II, we conclude that it is sufficient to discretize with the step equal to \( \tau = 1/1.028 = 1/1.028 = .028 \). We find that the set of matrices \( I + \tau A \) has a JSR smaller than .992 (computations have been done with the JSR toolbox \cite{33} and took less than a second on a standard desktop PC).

On the other hand, \( A + .2 I \) is unstable, and we conclude that

\[ \sigma(A) \in [-.2,0]. \]
III. PROOF OF THEOREM 3 USING THEOREM 4

We use the following Markov-Bernstein type inequality proved in [31]: for every $d \geq k + 1 \geq 1$ we have

$$|b_d^{(k)}(0)| \leq \frac{8^k(d-1)!k!}{(d-1-k)!(2k)!}. \hspace{1cm} (7)$$

Proof: (of Theorem 3). We have $S''(t) = e^{-t} \left( b''(t) - 2b'(t) + b(t) \right)$, hence $|S''(0)| \leq |b''(0)| + 2|b'(0)| + |b(0)|$. Substituting $|b(0)| = 1$ and estimating $|b''(0)|$ and $|b'(0)|$ by (7), we get $|S''(0)| \leq \frac{2}{3} (d-1)(2d-1) + 1$. Applying now Theorem 4, we arrive at Theorem 3.

IV. PROOF OF THEOREM 4

In this section we expose our technical developments that lead to our main result: Theorem 4. The proof will be split into three subsections. In the first part (Subsection IV-A) we estimate the value $s(a, \varepsilon)$ by a certain extremal problem on the set of quasipolynomials (polynomials of exponents). Then (Subsection IV-B) we describe its solution by the corresponding Tchebychev polynomial of exponents. Finally (Subsection IV-C), we estimate the numerical solution of the extremal problem.

A. An extremal problem on quasipolynomials

Let $h = (h_1, \ldots, h_d)$ be a positive vector. Denote by $\mathcal{P}_h$ the set of quasipolynomials $p(t) = \sum_{k=1}^d p_k e^{-h_k t}$. This is a $d$-dimensional subspace of the space $C_0(\mathbb{R}_+)$ of functions continuous on $\mathbb{R}_+$ and converging to zero as $t \to \infty$. For a given $\varepsilon > 0$, we denote by $\kappa(h, \varepsilon)$ the value of the following minimization problem:

$$\begin{align*}
\min & \quad \frac{1-p(0)}{p''(0) - \varepsilon p(0)}, \\
\text{subject to} & \quad p \in \mathcal{P}_h, \\
& \quad \|p\|_{C(\mathbb{R}_+)} \leq 1, \quad p'(0) > \varepsilon p(0).
\end{align*} \hspace{1cm} (8)$$

For a given matrix $B$ and for $x \in \mathbb{R}^d$, we denote $G_B(x) = \text{co}_h \{ e^{tB}x, \ t \in [0, +\infty) \}$, where $\text{co}_h(X) = \text{co} \{X, -X\}$ is the symmetrized convex hull of $X$. Thus, $G_B(x)$ is the convex hull of the curve $\gamma(t) = \{ e^{tB}x, \ t \in [0, +\infty) \}$ and of its reflection through the origin. If the matrix $A$ is Hurwitz, i.e., the real parts of all its eigenvalues are negative, then the set $G_B(x)$ is bounded, and the curve $\gamma$ connects the point $x = \gamma(0)$ with the origin $0 = \gamma(+\infty)$.

Proposition 1: For every matrix $B$ with a real negative spectrum, for every $x \in \mathbb{R}^d$ and $\varepsilon > 0$, the following holds: the largest number $\tau$ such that $x + \tau (B - \varepsilon I) x \in G_B(x)$ is equal to $\kappa(h, \varepsilon)$, where $h = -\varepsilon \text{sp}(B)$.

Proof: We assume without loss of generality that $B$ has $d$ distinct eigenvalues. (The assertion for general matrices follows by taking the limit.) By the Carathéodory theorem, a point belongs to the convex set $G_B(x)$ if and only if this point is a convex combination of at most $d + 1$ extreme points of that set. Each extreme point of the set $G_B(x)$ has the form $x + \varepsilon tBx$, $t \geq 0$. Hence, there are $n \leq d + 1$ nonnegative numbers $\{t_k\}_{k=1}^n$ and $n$ numbers $\{q_k\}_{k=1}^n$ such that

$$x + \tau (B - \varepsilon I) x = \sum_{k=1}^n q_k e^{-t_k B} x, \quad \sum_{k=1}^n |q_k| = 1.$$ \hspace{1cm} (9)

Now let us pass to the basis of eigenvectors of the matrix $B$. In this basis we denote $B = \text{diag}(-\beta_1, \ldots, -\beta_d)$, $\beta_i > 0$, $e^{tB} = \text{diag}(e^{-\beta_1 t}, \ldots, e^{-\beta_d t})$, and $x = (x_1, \ldots, x_d)$. We can assume that all coordinates of $x$ are nonzero; the assertion for general $x$ will again follow by taking the limit.

Writing (9) coordinatewise, we obtain

$$x_j \left( 1 - \tau (\beta_j + \varepsilon) \right) = \sum_{k=1}^n q_k e^{-t_k \beta_j} x_j,$$

or, eliminating $x_j$:

$$1 - \tau (\beta_j + \varepsilon) = \sum_{k=1}^n q_k e^{-t_k \beta_j}.$$ 

This equality does not involve $x$. Thus, the largest $\tau$ such that $x + \tau (B - \varepsilon I) x \in G_B(x)$ is the same for all $x \neq 0$.

Taking $x = e$ we observe that for every $\delta > 0$ the assertion $e + \delta (B - \varepsilon I)e \notin G_B(e)$ is equivalent to the existence of a linear functional $\tilde{p} = (p_1, \ldots, p_d) \in \mathbb{R}^d$ separating the point $e + \delta (B - \varepsilon I)e$ from the set $G_B(e)$, i.e.,

$$\tilde{p}, e + \delta (B - \varepsilon I)e > \max_{y \in G_B(e)} \tilde{p}, y.$$ 

The right hand side is equal to

$$\sup_{t \in \mathbb{R}_+} \left| \sum_{k=1}^d p_k e^{-t \beta_k} \right| = \|p\|_{C(\mathbb{R}_+)}.$$ 

where $p(t) = \sum_{k=1}^d p_k e^{-t \beta_k}$ is a quasipolynomial. On the other hand,

$$\tilde{p}, e + \delta (B - \varepsilon I)e = \sum_{k=1}^d p_k + \delta \sum_{k=1}^d p_k (-\beta_k - \varepsilon) = p(0) + \delta p'(0) - \delta \varepsilon p(0).$$

Normalizing, we get $\|p\|_{C(\mathbb{R}_+)} \leq 1$ and $p(0) + \delta p'(0) - \delta \varepsilon p(0) > 0$. Since $p(0) \leq \|p\|_{C(\mathbb{R}_+)} = 1$, we conclude that $p(0) > \varepsilon p(0)$, and hence $\delta > \frac{1-p(0)}{p'(0) - \varepsilon p(0)}$. Thus, the point $e + \delta (B - \varepsilon I)e$ does not belong to $G_B(e)$ if and only if $\delta > \frac{1-p(0)}{p'(0) - \varepsilon p(0)}$, which completes the proof.

Remark 1: In fact we have proved a bit more: for every $\tau < s(a, \varepsilon)$ and $x \neq 0$, the point $x + \tau (B - \varepsilon I)x$ is in the interior of $G_B(x)$.

Proposition 2: If all matrices from $A$ have real spectra, then

$$s(a, \varepsilon) \geq \min_{A \in \mathcal{A}} \kappa(h, \varepsilon), \hspace{1cm} (10)$$

where $h = -\varepsilon \text{sp}(A) - \varepsilon e$.

Proof: We need to show that if $\sigma(A) < -\varepsilon$, then $\rho(I + \tau A) < 1$, for all $\tau$ smaller than the right-hand side of (10).
It is well known (see [23], [24]) that \( \sigma(A) < -\varepsilon \) implies that there exists a norm in \( \mathbb{R}^d \) such that \( \|x(t)\| \leq e^{-\varepsilon t} \|x_0\| \), for every trajectory \( x(\cdot) \) with \( x_0 \neq 0 \).

In particular, this holds for a trajectory without switching, i.e., for a constant function \( A(t) \). Thus, for every \( A \in A \) we have \( \|e^{tA}x_0\| < e^{-\varepsilon t} \|x_0\| \), and hence \( \|e^{t(A+\varepsilon I)}x_0\| < \|x_0\| \) for all \( t \in \mathbb{R}_+ \).

Therefore, for any point \( y \) from the symmetrized convex hull of the set \( \{e^{t(A+\varepsilon I)}x_0 \mid t \in \mathbb{R}_+ \} \), we have \( \|y\| \leq \|x_0\| \).

Applying now Proposition 1 for the matrix \( B = A + \varepsilon I, \) and taking into account Remark 1, we see that \( (I + \tau A)x_0 \) is an interior point of the set \( G_{A+\varepsilon I}(x_0) \), and, consequently, \( \|(I + \tau A)x_0\| < \|x_0\| \). This means that the norm of the operator \( I + \tau A \) is smaller than 1, for each \( A \in A \). Whence, \( \rho(I + \tau A) < 1 \).

\[ B. \text{Tchebychev polynomials of exponents} \]

In view of Proposition 2, to estimate the value \( s(a, \varepsilon) \) from below it suffices to solve the extremal problem (8) with \( h = -sp(A) - \varepsilon e \) for each matrix \( A \in A \), and then take the smallest value among all the matrices. The solution of this problem is always unique and is readily available by the corresponding Tchebychev polynomial of exponents which possesses \( d \) points of alternance on \( \mathbb{R}_+ \). This polynomial can be constructed numerically by the Remez algorithm [6], [28].

In the next subsections we obtain estimates for this value in terms of the dimension \( d \) and the largest spectral radius of matrices from \( A \).

We start with several simple observations. First of all, for every \( \lambda > 0 \) we have \( \kappa(\lambda h, \lambda \varepsilon) = \lambda^{-1} \kappa(h, \varepsilon) \). Hence, everything can be reduced to the case when the largest exponent \( h_1 \) is 1. In the sequel we assume that \( 1 = h_d \geq \cdots \geq h_1 > 0 \). Furthermore, we allow some of the exponents \( h_d \) to coincide. In this case, if, for instance, \( h_1 = \cdots = h_m = h_{m+1} \), we say that \( h_1 \) has multiplicity \( m \) and the functions \( e^{-h_1 t}, e^{-h_2 t}, \ldots, e^{-h_d t} \) are replaced by \( e^{-h_1 t}, te^{-h_2 t}, \ldots, m^{-1}e^{-h_d t} \) respectively.

It is well known that the system of functions \( e^{-h_1 t}, \ldots, e^{-h_d t} \) (all the exponents may have multiplicities) is a Tchebychev system on the domain \( K = [0, +\infty) \), i.e., every polynomial (linear combination of several functions) on this system \( p \in \mathcal{P}_h \) has at most \( d - 1 \) zeros (see, for instance, [17], [19]). Let \( \mathcal{P} \) be the space of polynomials of some Tchebychev system of \( d \) elements on a domain \( K \).

For every \( \varepsilon > 0 \) and \( h = (h_1, \ldots, h_d) \) there is a unique polynomial \( p \in \mathcal{P} \) such that \( p(t_i) = e^{i\varepsilon}, i = 1, \ldots, d \). By Haar’s theorem [19], for every continuous function \( f \in C(K) \), there is a unique element \( p \in \mathcal{P} \) of best approximation, for which the value \( \|f - p\|_{C(K)} \) attains its minimum on the set \( \mathcal{P}_h \). By Karlin’s theorem (the “snake theorem,” [16]), the difference \( f - p \) is either identically zero, or it possesses \( l \) points of Tchebychev alternance, where \( f - p \) takes values equal by module with alternating signs. The number of points \( l \) of the alternance depends on \( K \) and on \( \mathcal{P} \). For the system \( e^{-h_1 t}, \ldots, e^{-h_d t} \) on \( K = \mathbb{R}_+ \), it follows that there exists a unique polynomial \( T = T_h \in \mathcal{P}_h \) and a unique system of points \( 0 = \nu_1 < \nu_2 < \cdots < \nu_d < \infty \) such that \( \|T\|_{C(\mathbb{R}_+)} = 1 \) and \( T(\nu_k) = (-1)^k, k = 1, \ldots, d \). We call it the \( h \)-Tchebychev polynomial. This is a polynomial from \( \mathcal{P}_h \) with the smallest deviation from zero among all polynomials with a given leading coefficient (i.e., coefficient for \( t^m e^{-h_1 t} \), where \( m \) is the multiplicity of \( h_d \)). This polynomial enjoys many extremal properties on the set \( \{p \in \mathcal{P}_h : \|p\|_{C(\mathbb{R}_+)} \leq 1 \} \). For example, for every \( t \leq 0 \) and for each nonnegative integer \( n \) the largest value of the \( n \)th derivative at the point \( t \) for this set of polynomials is attained on the polynomial \( T \), if \( n \) is odd, and \(-T, \) if \( n \) is even. (That is, \( T \) gives unique solutions to the problems \( p'(0) \to \max, \|p\|_{C(\mathbb{R}_+)} \leq 1 \)

\[ p''(0) \to \min, \|p\|_{C(\mathbb{R}_+)} \leq 1 \ldots \] Observe that the value of the first problem exceeds \( h_d \), because already for \( p = -e^{-h_1 t} \) we have \( p'(0) = h_d \).

Let \( \nu_2 \) be the smallest positive point of the alternance of the \( h \)-Tchebychev polynomial \( T \). Thus, \( T(\nu_2) = 1 \), and \( T(\cdot) \) is increasing and concave on the segment \( [0, \nu_2] \). For an arbitrary \( \varepsilon > 0 \) we consider the following problem:

\[ \begin{cases} 1 - T(t) & \to \min, \\ T'(t) - \varepsilon T(t) & \to \min, \\ t \in [0, \nu_2], & T'(t) > \varepsilon T(t). \end{cases} \]

In contrast to problem (8) this problem is easily solvable, just by finding a unique root of the derivative of the objective rational function \( 1 - T(t) / T'(t) - \varepsilon T(t) \) on the segment \( [0, \nu_2] \).

Proposition 3: For every \( \varepsilon > 0 \) and \( h = (h_1, \ldots, h_d) \) such that \( 0 < h_1 \leq \cdots \leq h_d = 1 \), the values of problems (8) and (11) coincide. This value is bigger than \( \frac{2 \varepsilon}{T''(0) + 2 \varepsilon} \), where \( T = T_h \).

The proof is in our forthcoming paper [27]. Proposition 3 allows us to obtain lower bounds for the value of problem (8) merely by estimating the value \( T''(0) \) for the corresponding \( h \)-Tchebychev polynomial \( T = T_h \), which leads, by means of Proposition 2, to a lower bound for \( s(a, \varepsilon) \). However, for most vectors \( h = (h_1, \ldots, h_d) \), the polynomial \( T \) cannot be found explicitly, but evaluated numerically with the Remez algorithm. Besides, it would be more convenient to have some uniform lower bounds for all values of \( h \). In the next section such bounds will be derived, which will complete the proof of Theorem 4.

C. The lower bound as a function of the dimension \( d \) and the maximal spectral radius \( r(\mathcal{A}) \)

We start with the following key result, whose full proof is to be found in our forthcoming paper [27].

**Lemma 1:** Suppose \( h_1 \) has multiplicity \( n \leq d - 1 \); then, for arbitrary \( \varepsilon > 0 \) and for all sufficiently small \( \delta > 0 \) the value of problem (8) decreases after replacement of all \( h_1, \ldots, h_n \) by \( h_1 + \delta, \ldots, h_n + \delta \).

Thus, one can always slightly increase the smallest exponent \( h_1 \) to reduce the value of problem (8). On the other hand, the set of quasipolynomials \( p \) of \( d \) terms such that \( \|p\|_{C(\mathbb{R}_+)} \leq 1 \), and whose exponents are in the segment \([h_1, h_d]\) is compact. Hence, the minimal value of problem (8)
is attained at the polynomial with $d$ exponents $h_d$, i.e., for a polynomial $p(t) = e^{-h_d t} \sum_{k=0}^{d-1} p_k t^k$. By Proposition 3 this extremal polynomial has $d$ points of alternance on $\mathbb{R}_+$, i.e., coincides with $T_h$, where $h = (h_d, \ldots, h_d)$. Moreover, if we change the variable $t' = h_d t$, we reduce the whole analysis to just one polynomial with $h_d = 1$. This is nothing else as the L-Tchebychev polynomial $S_d(t) = e^{-t} b_d(t)$.

Theorem 4 can now be proved by combining the above results. The full proof is to be found in the journal version of this work [27].

V. CONCLUSION

The goal of this paper was to provide a way to compute the maximal rate of growth of a trajectory of a continuous time switching linear system, with a bound on the computation time necessary to do it with a specified accuracy. We showed that this is possible for matrices with real spectrum, and we leave open the question for general matrices. Our techniques have a different flavor than the ones previously proposed in the literature, as they mainly aim at computing a discretization step, in order to apply efficient methods for discrete time systems (like the ones implemented in the toolbox [33]). Our theorems are surprisingly technical, especially compared with the simplicity of their main product, that is, the recipe (4) to compute the discretization step $\tau$. We believe that our main results can be extended to general matrices, without the real spectrum assumption. However, this, probably, requires a different technique. Indeed, our approach uses essentially that any finite collection of real exponents constitute a Tchebychev system, which is not the case for complex exponents.

VI. ACKNOWLEDGEMENTS

We thank the anonymous referees for their helpful comments. The research reported here has been carried out while V.P. was visiting UCLouvain. That author wishes to thank the University for its hospitality.

REFERENCES