A MAXIMUM PRINCIPLE FOR THE STABILITY ANALYSIS OF
POSITIVE BILINEAR CONTROL SYSTEMS WITH APPLICATIONS
TO POSITIVE LINEAR SWITCHED SYSTEMS

LIOR FAINSHIL AND MICHAEL MARGALIOT∗

Abstract. We consider a continuous-time bilinear control system with Metzler matrices. Each entry in the transition matrix of such a system is non-negative, making the positive orthant an invariant set of the dynamics. Motivated by the stability analysis of positive linear switched systems (PLSSs), we define a control as optimal if, for a fixed final time, it maximizes the spectral radius of the transition matrix. Our main result is a first-order necessary condition for optimality in the form of a maximum principle (MP). The proof of this MP combines the standard needle variation with a basic result from the Perron-Frobenius theory of non-negative matrices. We describe several applications of this MP to the stability analysis of PLSSs under arbitrary switching.

Key words. Positive linear switched systems, stability under arbitrary switching law, positive linear systems, Metzler matrix, variational approach, Perron-Frobenius theory, positive absolute stability problem.

1. Introduction. Two matrices \( A_0, A_1 \in \mathbb{R}^{n \times n} \) give rise to the continuous-time linear switched system:

\[
\begin{align*}
\dot{x}(t) &= A_{\sigma(t)}x(t), \\
x(0) &= x_0,
\end{align*}
\]

where \( x : \mathbb{R}_+ \to \mathbb{R}^n \) is the state vector, and \( \sigma : \mathbb{R}_+ \to \{0, 1\} \) is a piecewise constant function referred to as the switching signal. This models a system that can switch between the two linear subsystems:

\( \dot{x} = A_0x \) and \( \dot{x} = A_1x. \)

We say that (1.1) is globally uniformly asymptotically stable (GUAS) if there exists a class K function\(^1\) \( \beta \) such that for any initial condition \( x_0 \in \mathbb{R}^n \) and any switching law \( \sigma \), the corresponding solution of (1.1) satisfies

\[
|x(t)| \leq \beta(|x_0|, t), \quad \text{for all } t \geq 0.
\]

This implies in particular that

\[
\lim_{t \to \infty} x(t) = 0 \quad \text{for any switching law } \sigma \text{ and any } x_0 \in \mathbb{R}^n.
\]

For linear switched systems, (1.2) is equivalent to GUAS (see, e.g., [4]).

It is well-known and easy to demonstrate that the following assumption is a necessary (but not sufficient) condition for GUAS of (1.1).

Assumption 1 For any \( k \in [0, 1] \), the matrix \( kA_0 + (1 - k)A_1 \) is Hurwitz.

\(^{1}\)A continuous function \( \alpha : [0, \infty) \to [0, \infty) \) belongs to the class K if it is strictly increasing and \( \alpha(0) = 0 \). A continuous function \( \beta : [0, \infty) \times [0, \infty) \to [0, \infty) \) belongs to the class K if for each fixed \( s \), \( \beta(s, \cdot) \) belongs to \( K \), and for each fixed \( r > 0 \), the mapping \( \beta(r, \cdot) \) is decreasing and \( \beta(r, s) \to 0 \) as \( s \to \infty \).
There is a rich literature on sufficient conditions for GUAS, see, e.g., [14; 16; 33; 34; 58]. A more challenging problem is to determine a necessary and sufficient condition for GUAS. What makes this problem difficult is that the set of all possible switching laws is huge, so exhaustively checking the solution for each switching law is impossible.

A natural idea is to try and characterize the “most destabilizing” switching law $\sigma^*$, and then analyze the behavior of the corresponding trajectory $x^*$. If $x^*$ converges to the origin, then so does any trajectory of the switched system and this establishes GUAS. This idea was pioneered by E. S. Pyatntisky [51; 52], who studied the celebrated absolute stability problem that is equivalent to finding a necessary and sufficient condition for GUAS of (1.1) under the additional condition

$$\text{rank}(A_1 - A_0) = 1,$$

i.e. $A_1 = A_0 + bc'$ for some $b, c \in \mathbb{R}^n$. Pyatntisky proposed an optimal control approach for studying this problem. This variational approach was further developed by several scholars including and N. E. Barabanov and L. B. Rapoport, and proved to be highly successful; see the survey papers [37; 10; 53], the related work in [6; 7], and the recent extensions to the stability analysis of discrete–time linear switched systems in [48; 49].

Recall that a linear system $\dot{x} = Ax$, with $A \in \mathbb{R}^{n \times n}$, is called positive if the positive orthant $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n \mid x_i \geq 0, i = 1, \ldots, n\}$ is an invariant set of the dynamics, that is, if $x(0) \in \mathbb{R}^n_+$ implies that $x(t) \in \mathbb{R}^n_+$ for all $t \geq 0$. A necessary and sufficient condition for this is that $A$ is a Metzler matrix, that is, $a_{ij} \geq 0$ for any $i \neq j$. This implies that for any $t \geq 0$ every entry of $\exp(At)$ is non–negative. By the Perron–Frobenius theory, the eigenvalue of $\exp(At)$ with maximal absolute value is real and positive.

Positive linear systems play an important role in systems and control theory because in many physical systems the state-variables represent quantities that can never attain negative values (e.g. population sizes or protein concentrations) [19; 11; 30].

If both $A_0$ and $A_1$ are Metzler and $x_0 \in \mathbb{R}^n_+$, then we refer to (1.1) as a positive linear switched system (PLSS). PLSSs were used for modeling communication systems [56] and formation flying [27] (see also [57; 22; 44]). Mason and Shorten [45], and independently David Angeli, posed the following conjecture.

**Conjecture 1** If (1.1) is a PLSS, then Assumption 1 provides a sufficient condition for GUAS.

Recently, Gurvits, Shorten, and Mason [24] proved that this conjecture is in general false (see also [23]). However, they also showed that it does hold when $n = 2$ (even when the number of matrices is arbitrary). Their proof in the planar case is based on showing that the system admits a common quadratic Lyapunov function. (For more on the analysis of switched systems using common Lyapunov functions, see [14; 57; 44; 22].) Margaliot and Branicky [39] used the variational approach to derive a reachability–with–nice–controls–type result for planar bilinear control systems, and showed that the proof of Conjecture 1 when $n = 2$ follows as a special case. Fainshil, Margaliot, and Chigansky [18] showed, using a specific example, that
Conjecture 1 is false already for the case $n = 3$. These results suggest that as far as the GUAS problem is concerned, analyzing PLSSs is, in general, not simpler than analyzing linear switched systems.

In this paper, we develop a variational approach for analyzing PLSSs. The first step is to embed the PLSS in a more general positive bilinear control system. The transition matrix $C(t)$ of such a system is nonnegative for any $t \geq 0$. For a fixed final time $T > 0$, we pose the problem of finding a control that maximizes the spectral radius of $C(T)$. Our main result is a new MP that provides a necessary condition for optimality. The proof of this MP combines the standard idea of introducing a needle variation with a basic result from the Perron-Frobenius theory. We describe several applications of the new MP to the stability analysis of PLSSs.

We note that a first step in this direction was taken in [40]. However the approach in that paper uses the classical Pontryagin maximum principle (PMP).

The remainder of this paper is organized as follows: Section 2 reviews the problem of absolute stability and its variant of positive absolute stability. Section 3 details some known results from the Perron-Frobenius theory of non-negative matrices that are used later on. Section 4 relates the stability of a positive bilinear control system to the generalized spectral radius of its transition matrix. This motivates the optimal control problem of maximizing the spectral radius of the transition matrix. Section 5 introduces our main result, namely, a first-order maximum principle for this optimal control problem. Several implications of this MP are described in Section 6. Applications to the stability analysis of PLSSs are described in Section 7.

We use standard notation. Matrices are denoted by a capital letter. $V'$ denotes the transpose of the matrix $V$. Column vectors are denoted by small letters (e.g. $x$), so $x'$ is a row vector. For a matrix $V \in \mathbb{R}^{n \times k}$, the notation $V > 0$ ($V \geq 0$) means that every entry of $V$ is positive (non-negative).

2. Absolute stability. We begin by reviewing the variational approach derived in the context of the celebrated absolute stability problem (ASP). We also describe a variant of the ASP called the positive absolute stability problem (PASP). Solving the ASP [PASP] amounts to finding a necessary and sufficient condition for GUAS of a positive linear switched system that satisfies (1.3).

Consider the system

$$\dot{x}(t) = Ax(t) + b\phi(t, c'x(t)), \quad (2.1)$$

where $A$ is Hurwitz, the pair $(A, b)$ [(A, c)] is controllable [observable], and $\phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ belongs to $S_k$, the set of scalar time-varying functions in the sector $[0, k]$, i.e. $\phi(t, 0) = 0$ and $0 \leq \phi(t, z) \leq k z^2$ for all $t, z$. For example, every time-invariant linear function $\phi(z) = cz$, with $c \in [0, k]$, belongs to $S_k$.

As depicted in Fig. 2.1, we can view (2.1) as the feedback connection of the SISO linear control system

$$\dot{x} = Ax + bu,$$  

$$y = c'x, \quad (2.2)$$

and a function $\phi \in S_k$.

By definition, $S_0$ contains only the function $\phi(t, z) \equiv 0$. Since $A$ is Hurwitz, (2.1) is asymptotically stable when $\phi \in S_0$. By continuity, there exists $\epsilon > 0$ such that (2.1) is asymptotically stable for any $\phi \in S_\epsilon$. This naturally leads to the following problem.
Problem 1 (Absolute Stability) Find the value:

\[ k^* := \min \{k > 0 : \text{there exists } \phi \in S_k \text{ such that (2.1) is not asymptotically stable} \} \]

In other words, for \( k \in [0, k^*) \), (2.1) is asymptotically stable for any \( \phi \in S_k \), and there exists a function \( \phi^* \in S_{k^*} \), referred to as a worst-case nonlinearity, for which (2.1) is not asymptotically stable.

This celebrated problem, dating back to the 1940s [35; 54; 59], led to numerous important results in the mathematical theory of stability and control, including: Popov’s criterion; the circle criterion; the positive-real lemma [13]; and the theory of integral quadratic constraints [47]. An analysis of the computational complexity of some closely related problems can be found in [12].

Since \( \phi \in S_k \), \( \phi(t, c'x(t)) = h(t, c'x(t))c'x(t) \) with \( 0 \leq h(t, c'x(t)) \leq k \) for all \( t \). Let \( B_k = kbc' \), and consider the bilinear control system (BCS):

\[ \dot{x} = (A + B_k u)x, \quad u \in U, \quad (2.3) \]

where \( U \) is the set of measurable functions taking values in \([0, 1]\). We say that (2.3) is globally asymptotically stable (GAS) if \( \lim_{t \to \infty} x(t) = 0 \) for any control \( u \in U \) and any initial condition \( x(0) \in \mathbb{R}^n \).

It is clear that (2.3) is GAS for \( k = 0 \). It can be shown that

\[ k^* = \min \{k > 0 : \text{system (2.3) is not GAS} \}. \]

In other words, solving the ASP is equivalent to finding a necessary and sufficient condition for GAS of (2.3). Note that (2.3) is a bilinear control system [17] with the special property that the control term \( B_kux \) includes a rank one matrix, as \( \text{rank}(B_k) = \text{rank}(kbc') = 1 \).

In the 1940s, M. A. Aizerman posed the following conjecture.

Conjecture 2 Let \( s^* \in (0, \infty] \) be the minimal value \( s > 0 \) such that \( A + sbc' \) is not Hurwitz. Then \( k^* = s^* \).

In other words, the conjecture asserts that we may replace the time-varying nonlinearity with a linear element with a variable slope and then proceed as in linear analysis.
Paraphrased in terms of linear switched systems, the conjecture is that Assumption 1, which is a necessary condition for GUAS, is also a sufficient condition. This conjecture is false. In 1957, R. E. Kalman conjectured that a stronger condition implies absolute stability, but his conjecture is also known to be false [8].

E. S. Pyatnitsky [51; 52] pioneered a variational approach to tackle the ASP. This approach is motivated by an attempt to characterize a “most destabilizing” nonlinearity \( \phi^* \) (see the survey paper [37]). Paraphrased in terms of the BCS (2.3), the variational approach seeks to determine, using tools from optimal control theory, a “most destabilizing” control \( u^* \). If the trajectory corresponding to \( u^* \) converges to the origin, the BCS is GAS. This reduces the problem of analyzing stability for all \( u \in \mathcal{U} \) to the analysis of the system when \( u = u^* \).

The variational approach has several advantages. First, it allows the application of sophisticated and powerful tools, such as first– and higher–order maximum principles (MPs) to stability analysis. Second, some of the results can be generalized to nonlinear control systems and nonlinear switched systems. Third, it allows the derivation of not only stability results, but more general nice–reachability–type results [42; 39; 55] (see also [38] for some related considerations).

The variational approach was used to derive the most general stability results currently available for both: (1) linear switched systems of order \( n = 2 \) [52; 41], and \( n = 3 \) [9; 43] (see also the extensions to homogeneous switched systems in [21; 25]); and (2) nonlinear switched systems with a nilpotent Lie algebra [55].

If we assume that \( A \) and \( A + kbc' \) are Metzler and that \( x(0) \in \mathbb{R}^n_+ \), then any solution of (2.1) satisfies \( x(t) \in \mathbb{R}^n_+ \) for any \( t \geq 0 \). In this case, Problem 1 is referred to as the positive absolute stability problem.

3. Some results from the Perron–Frobenius theory. In this section we review some known results on non–negative matrices that will be used later on. For more details and proofs, see e.g. [11; 26].

Recall that a matrix \( P \in \{0, 1\}^{n \times n} \) is called a permutation matrix if exactly one entry in each row and column is equal to 1, and all other entries are 0. Multiplication by such a matrix effects a permutation of the rows or columns of the object multiplied.

**Definition 1** A matrix \( A \in \mathbb{R}^{n \times n} \) is said to be reducible if either
1. \( n = 1 \) and \( A = 0 \); or
2. \( n \geq 2 \), there is a permutation matrix \( P \in \{0, 1\}^{n \times n} \), and an integer \( r \) with \( 1 \leq r \leq n - 1 \) such that \( P'AP = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} \), where \( B \in \mathbb{R}^{r \times r} \), \( D \in \mathbb{R}^{(n-r) \times (n-r)} \), \( C \in \mathbb{R}^{r \times (n-r)} \) and \( 0 \in \mathbb{R}^{(n-r) \times r} \) is a zero matrix.

A matrix \( A \in \mathbb{R}^{n \times n} \) is said to be irreducible if it is not reducible.

The spectral radius of a matrix \( A \) is \( \rho(A) = \max \{ |\lambda| : \lambda \text{ is an eigenvalue of } A \} \).

**Theorem 1** [26, Ch. 8] Suppose that \( A \in \mathbb{R}^{n \times n} \) is non-negative and irreducible. Then
1. \( \rho(A) \geq 0 \);
2. \( \rho(A) \) is an eigenvalue of \( A \);
3. \( \rho(A) \) is an algebraically (and hence geometrically) simple eigenvalue of \( A \);
4. There exists a unique (up to scaling) vector \( x > 0 \) such that \( Ax = \rho(A) x \).

**Definition 2** A nonnegative matrix \( A \in \mathbb{R}^{n \times n} \) is said to be primitive if it is irreducible and has only one eigenvalue of maximum modulus.
Theorem 2 [26, Ch. 8] Suppose that $A \in \mathbb{R}^{n \times n}$ is nonnegative and primitive. Let $\tilde{A} = \frac{A}{\rho(A)}$, then

$$\lim_{m \to \infty} (\tilde{A})^m = xy',$$

where $x > 0$ is the eigenvector of $A$ corresponding to $\rho(A)$, $y > 0$ is the eigenvector of $A'$ corresponding to $\rho(A)$, and $x'y = 1$.

We refer to $x, y$ as the dominant eigenvectors. To provide an intuitive explanation of (3.1), assume that $A$ is diagonalizable, i.e. $T^{-1}AT = D = \text{diag}(\lambda_1, \ldots, \lambda_n)$, with the eigenvalues arranged so that $\lambda_1 = \rho(A)$. Then $\tilde{A} = \frac{A}{\lambda_1} = T \frac{D}{\lambda_1} T^{-1}$, so

$$(\tilde{A})^m = T(D/\lambda_1)^m T^{-1} = T \text{diag}(1, (\lambda_2/\lambda_1)^m, \ldots, (\lambda_n/\lambda_1)^m) T^{-1}.$$ 

Letting $e^1$ denote the first column of the $n \times n$ identity matrix, this implies that $\lim_{m \to \infty} (\tilde{A})^m = xe^1$, with $x = Te^1$, $y = (T')^{-1}e^1$. Note that $x'y = 1$. Since $T^{-1}ATe^1 = De^1 = \lambda_1 e^1$, we have $Ax = \lambda_1 x$, so $x$ is an eigenvector of $A$ corresponding to $\lambda_1$. Similarly, $y$ is an eigenvector of $A'$ corresponding to $\lambda_1$.

In the next section, we define the generalized spectral radius of a positive bilinear control system and relate it to the GAS property. This provides the motivation for the optimal control problem posed in Section 5.

4. Generalized spectral radius of a positive bilinear control system.

The first step in the variational approach is relaxing the linear switched system (1.1) to a BCS

$$\dot{x} = (A + uB)x, \quad u \in \mathcal{U},$$

$$x(0) = x_0,$$

with $A = A_0$ and $B = A_1 - A_0$. Recall that the BCS is said to be GAS if $\lim_{t \to \infty} x(t) = 0$ for any initial condition $x_0$ and any control $u \in \mathcal{U}$. Note that every trajectory of the switched system (1.1) is also a trajectory of (4.1), so GAS of (4.1) implies GUAS of the linear switched system. It is not difficult to show that the converse implication also holds, so the BCS is GAS if and only if the linear switched system is GUAS.

From here on we assume that $A + kB$ is Metzler for any $k \in [0, 1]$. This implies that if $x_0 \in \mathbb{R}^n_+$, then $x(t) \in \mathbb{R}^n_+$ for any $u \in \mathcal{U}$ and any $t \geq 0$. We refer to (4.1) as a positive bilinear control system (PBCS).

For $0 \leq a \leq b \leq T$, let $C(b, a)$ denote the value at time $t = b$ of an absolutely continuous solution of

$$\frac{d}{dt} C(t, a) = (A + Bu(t))C(t, a),$$

$$C(a, a) = I.$$  (4.2)

It is straightforward to verify that for any $u \in \mathcal{U}$ and any $0 \leq a \leq b \leq T$, the solution of (4.1) satisfies $x(b) = C(b, a)x(a)$. In other words, $C(b, a)$ is the transition matrix from time $a$ to time $b$ of (4.1) corresponding to the control $u$.

By differentiating the identity

$$C(b, a)C(a, 0) = C(b, 0).$$
with respect to $a$ and using (4.2), we find that for almost all $a \in [0,T]$,

$$
\frac{d}{da} C(b,a) = -C(b,a)(A + Bu(a)).
$$

(4.3)

We will be mainly interested in the case where $a = 0$ for which we write (4.2) as

$$
\dot{C}(t) = (A + Bu(t))C(t),
$$

$C(0) = I.$

(4.4)

For a PBCS, $C(t)$ is a non-negative matrix for any $t \geq 0$ and any $u \in U$. Since $C(t)$ is also non-singular, $\rho(C(t))$ is a real and positive eigenvalue of $C(t)$.

To relate $\rho(C(t))$ to the stability of the PBCS, we define the generalized spectral radius of (4.1) by

$$
\rho(A, B) = \limsup_{t \to \infty} \rho_t(A, B),
$$

where

$$
\rho_t(A, B) = \max_{u \in U} \rho(C(t))^{1/t},
$$

(4.5)

Note that the maximum here is well-defined, as the reachable set of (4.4) corresponding to $U$ is closed and bounded [20]. In fact, this is why we consider a control system with controls in $U$ rather than the original linear switched system.

The next example provides some intuition on $\rho(A, B)$ by considering the trivial case where the PBCS reduces to a positive linear system.

**Example 1** Suppose that $B = 0$, so that (4.1) becomes the positive linear system $\dot{x} = Ax$. In this case, $C(t) = \exp(At)$ for any $u$, so $(\rho_t(A, B))^t = \rho(\exp(At))$. Let $\lambda_i = \alpha_i + j\beta_i$, $i = 1, \ldots, n$, denote the eigenvalues of $A$, ordered so that $\alpha_1 \geq \alpha_i$ for all $i$. Then,

$$
(\rho_t(A, B))^t = \max_i \{ |\exp(\lambda_i t)| \} = \max_i \{ |\exp(\alpha_i t)| \} = \exp(\alpha_1 t).
$$

Hence, $\rho(A, B) < 1$ if and only if $\alpha_1 < 0$, that is, if and only if $A$ is Hurwitz.

The next result relates the GAS of the PBCS and $\rho(A, B)$.

**Theorem 3** The PBCS (4.1) is GAS if and only if $\rho(A, B) < 1$.

The proof of this result is similar to the proof of an analogous result relating the stability of a discrete–time linear switched system to its generalized spectral radius; see [28, Ch. 2].

**Remark 1** It follows from Thm. 3 that if $\rho(C(T)) \geq 1$ for some $T > 0$ and $u \in U$, then the PBCS is not GAS. Indeed, for any integer $k > 0$, define $\bar{u} : [0,kT] \to [0,1]$ via the periodic extension of $u$, and let $\bar{C}(t)$ denote the corresponding solution of (4.4) at time $t$. Then

$$
\rho(\bar{C}(kT)) = (\rho(C(T)))^k,
$$
so (4.5) yields

\[
\rho_{kT}(A, B) \geq (\rho(C(kT)))^{1/(kT)} = (\rho(C(T)))^{1/T} \geq 1,
\]

and this implies that \( \rho(A, B) \geq 1 \).

Example 2 Consider the PBCS (4.1) with \( n = 3 \), \( A = \begin{bmatrix} -1 & 0 & 0 \\ 10 & -1 & 0 \\ 0 & 0 & -10 \end{bmatrix} \), and \( B = \begin{bmatrix} -9 & 0 & 10 \\ -10 & -9 & 0 \\ 0 & 10 & 9 \end{bmatrix} \). It is straightforward to verify that \( A + kB \) is Hurwitz and Metzler for any \( k \in [0, 1] \). We used a simple numerical algorithm to (approximately) calculate \( \max_{u \in U} \rho(C(t)) \) and \( \rho_t(A, B) \) for various values of \( t \). The results are depicted in Figs. 4.1 and 4.2. It may be seen that in this case \( \rho_t(A, B) \) converges to a limit \( l \) with \( l > 2.5 \), so the PBCS is not GAS.

The next section presents our main result.

5. Maximum principle for positive bilinear control systems. Thm. 3 motivates the following optimal control problem.

Problem 2 Consider the PBCS (4.4). Fix an arbitrary \( T > 0 \). Find a control \( u^* \in U \) that maximizes \( \rho(C(T)) \).

Our main result is a necessary condition for optimality, stated in the form of a maximum principle (MP).
Consider the PBCS (4.4). Suppose that $u^* \in \mathcal{U}$ is an optimal control for Problem 2. Let $C^*(t)$ denote the corresponding solution of (4.4) at time $t$, and let $p^* = p(C^*(T))$. Suppose that $C^*(T)$ is primitive. By Theorem 2, there exist $v^*, w^* \in \mathbb{R}^n_+$, such that $(v^*)'w^* = 1$ and $\lim_{k \to \infty} (\frac{C^*(T)}{p^*})^k = v^*(w^*)'$. Define $q : [0, T] \to \mathbb{R}^n_+$ by

$$\dot{q} = -(A + Bu^*)'q,$$

$$q(T) = w^*, \quad (5.1)$$

and $p : [0, T] \to \mathbb{R}^n_+$ by

$$\dot{p} = (A + Bu^*)p,$$

$$p(0) = v^*. \quad (5.2)$$

Define the switching function $m : [0, T] \to \mathbb{R}$ by

$$m(t) = q'(t)Bp(t). \quad (5.3)$$

Then for almost all $t \in [0, T]$,

$$u^*(t) = \begin{cases} 1, & m(t) > 0, \\ 0, & m(t) < 0. \end{cases} \quad (5.4)$$

Intuitively speaking, as $u^*(t) \in [0, 1]$ for all $t$ and $u^*$ is an optimal control, we may expect that there are times where it takes the extremal values 0 or 1. Theorem 4 describes such times, namely, for any time $t$ such that $m(t) > 0$ [or $m(t) < 0$] the optimal control must satisfy $u^*(t) = 1$ [or $u^*(t) = 0$].

**Remark 2** Note that since we are considering an optimization problem for the matrix differential equation (4.4), we may expect an MP that includes an adjoint described by a matrix differential equation as well (see e.g., [5]). However, the MP actually entails two vector differential equations. The (generally unknown) dominant eigenvectors $v^*, w^*$ provide the boundary conditions for these vector differential equations.
Remark 3 An important property of this MP is that the switching function satisfies \( m(T) = m(0) \). Indeed, by (5.2),
\[
p(T) = C^*(T)p(0) \\
= C^*(T)v^* \\
= \rho^*v^* \\
= \rho^*p(0).
\]
(5.5)

As for \( q \), it is straightforward to show using (4.3) that for any \( t \in [0, T] \),
\[
q'(t) = q'(T)C^*(T, t),
\]
(5.6)
where \( C^*(T, t) \) is the solution of (4.2) corresponding to \( u^* \). For \( t = 0 \) this yields
\[
q(0) = (C^*(T))^tq(T) \\
= (C^*(T))^t w^* \\
= \rho^*w^* \\
= \rho^*q(T).
\]
(5.7)
Combining this with (5.3) yields
\[
m(T) = q'(T)Bp(T) \\
= q'(0)Bp(0) \\
= m(0).
\]
(5.8)

Remark 4 As stated, (5.1) and (5.2) lead to a two-point boundary value problem. However, since \( \rho^* > 0 \) and since (5.4) only depends on the sign of \( m(t) \), it follows from (5.7) that we can replace \( q(T) = w^* \) in (5.1) by \( q(0) = w^* \), thus obtaining a one-point boundary value problem in the MP.

Proof of Theorem 4. Let \( u^* \in \mathcal{U} \) be an optimal control. Fix an arbitrary time \( \tau \in (0, T) \), and an arbitrary \( a \in [0, 1] \). Define a needle variation of the optimal control \( u^* \) by:
\[
u(t) = \begin{cases} 
a, & t \in [\tau, \tau + \epsilon), \\
u^*(t), & \text{otherwise},
\end{cases}
\]
with \( \epsilon > 0 \) (below we will consider the case where \( \epsilon \to 0 \)). Note that \( u \in \mathcal{U} \). Let \( C(t) \) denote the solution of (4.4) corresponding to \( u \) at time \( t \). (Obviously, \( C(\cdot) \) depends on the parameters \( a, \tau, \epsilon \) but for simplicity we use a notation that suppresses this dependence.) Then
\[
C(T) = C^*(T, \tau + \epsilon) \exp((A + aB)\epsilon)C^*(\tau, 0).
\]
(5.9)
Differentiating (5.9) with respect to \( \epsilon \) and using (4.3) yields
\[
\frac{d}{d\epsilon}C(T) = \left( \frac{d}{d\epsilon}C^*(T, \tau + \epsilon) + C^*(T, \tau + \epsilon)(A + aB) \right) \exp((A + aB)\epsilon)C^*(\tau, 0) \\
= C^*(T, \tau + \epsilon) (-A - Bu^*(\tau + \epsilon) + A + aB) \exp((A + aB)\epsilon)C^*(\tau, 0) \\
= (a - u^*(\tau + \epsilon))C^*(T, \tau + \epsilon)B \exp((A + aB)\epsilon)C^*(\tau, 0).
\]
(5.10)
Define \( \gamma : \mathbb{R}_+ \to \mathbb{R}_+ \) by \( \gamma(\epsilon) = \rho(C(T)) \). Note that \( \gamma(0) = \rho(C^*(T)) = \rho^* \).

Since \( \rho^* \) is a simple eigenvalue of \( C^*(T) \), it follows from known results (see, e.g., [36, Thm. 2]) that for any sufficiently small \( \epsilon > 0 \),

\[
\gamma(\epsilon) = \gamma(0) + \epsilon \dot{\gamma}(0) + o(\epsilon),
\]

where \( \dot{\gamma}(0) = (u^*)'(C(T)|_{\epsilon=0})v^* \) and \( o(\epsilon) \) denotes a function satisfying \( \lim_{\epsilon \to 0} \frac{o(\epsilon)}{\epsilon} = 0 \). If \( \dot{\gamma}(0) > 0 \), then this contradicts the optimality of \( u^* \). Thus,

\[
(w^*)' \left( \frac{d}{\epsilon} C(T)|_{\epsilon=0} \right) v^* \leq 0,
\]

and combining this with (5.10) yields

\[
(a - u^*(\tau))(w^*)' C^*(T, \tau) BC^*(\tau, 0)v^* \leq 0.
\]

Using the definitions of \( p, q, \) and (5.6) yields

\[
0 \geq (a - u^*(\tau))q'(\tau)C^*(T, \tau) BC^*(\tau, 0)p(0)
= (a - u^*(\tau))q'(\tau)Bp(\tau)
= (a - u^*(\tau))m(\tau).
\]

(5.11)

Suppose that \( m(\tau) > 0 \). Then \( a - u^*(\tau) \leq 0 \) and since \( a \) can take any value in \([0, 1]\), \( u^*(\tau) = 1 \). Similarly, \( m(\tau) < 0 \) implies that \( u^*(\tau) = 0 \). Since \( \tau \) is arbitrary, this proves (5.4).

The MP in Thm. 4 is of course implicit, as \( v^* \), \( w^* \) are generally unknown. We will see below that it may still be used to obtain interesting and explicit results. Before that we consider an example where we first numerically find \( u^* \), and then demonstrate that \( u^* \) indeed satisfies the necessary condition for optimality (5.4).

**Example 3** Consider the PBCS in Example 2. For \( T = 2 \), a numerical search suggests that

\[
C^*(T) = \exp((A + B)t_4) \exp(At_3) \exp((A + B)t_2) \exp(At_1),
\]

with \( t_1 = t_3 = 0.6632382, t_2 = t_4 = 0.33676183 \). A calculation yields \( \rho(C^*(T)) = 6.0641 \), and

\[
v^* = \begin{bmatrix} 0.68729212 & 0.03311261 & 0.72562605 \end{bmatrix}' \]

\[
w^* = \begin{bmatrix} 1.44395608 & 0.21614296 & 0.00058337 \end{bmatrix}'.
\]

(5.13)

We numerically solved the differential equations in Thm. 4 with the boundary values determined by (5.13), and calculated \( m(t) \) and \( w^*(t) \). The switching function \( m(t) \) is shown in Fig. 5.1. Note that the zeros of \( m(t) \) in \([0, T]\) are \( 0, t_1, t_1 + t_2, t_1 + t_2 + t_3 \), and this implies that \( u^*(t) \) in (5.4) indeed yields the solution in (5.12).

**5.1. Hamiltonian Form.** As expected, Thm. 4 may be expressed in a Hamiltonian form. Define \( H : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R} \) by \( H(p, q, v) = q'(A + vB)p \). Then (5.1), (5.2), and (5.4) may be written as

\[
\dot{q}(t) = -\frac{\partial H}{\partial p}(p(t), q(t), u^*(t)),
\]
\[ \dot{p}(t) = \frac{\partial H}{\partial q}(p(t), q(t), u^*(t)), \]

and

\[ u^*(t) = \arg \max_{v \in [0, 1]} H(p(t), q(t), v), \]

respectively. A standard argument (see e.g. [31, Ch. 4]) can be used to show that

\[ H(t) = H(p(t), q(t), u^*(t)) \]

is constant, i.e., \( H(t) = H(0) \) for all \( t \in [0, T] \).

### 5.2. Optimal final time

Theorem 4 considers the case of an arbitrary final time \( T > 0 \). We now consider the case where \( T \) itself is optimal in some (local) sense.

**Definition 3** We say that \( T > 0 \) is **locally maximal** [minimal] if there exists \( \epsilon > 0 \) such that \( \rho(C^*(T)) \geq \rho(C^*(T+s)) \) [\( \rho(C^*(T)) \leq \rho(C^*(T+s)) \)] for any \( s \in (-\epsilon, \epsilon) \).

We say that \( T > 0 \) is **extremal** if it is either locally maximal or locally minimal.

It is important to note that \( C^*(T) \) and \( C^*(T+s) \) may correspond to different optimal controls, as the optimal control depends on the final time.

Using the standard tool of a *temporal control perturbation* (see, e.g. [32, Ch. 4]), it is possible to strengthen the MP when \( T \) is extremal.

**Theorem 5** Suppose that the conditions of Theorem 4 hold and that the final time \( T > 0 \) is extremal. Then the constant value of the Hamiltonian \( H(t) = H(p(t), q(t), u^*(t)) \) is zero.

**Proof.** We consider the case where \( T \) is locally maximal (the proof when it is locally minimal is similar). For \( \epsilon > 0 \) sufficiently small, define a control \( u : [0, T + \epsilon] \to [0, 1] \) by

\[ u(t) = \begin{cases} 
  u^*(t), & t \in [0, T), \\
  u^*(T), & t \in [T, T + \epsilon]. 
\end{cases} \]
The corresponding solution satisfies
\[ C(T + \epsilon) = \exp((A + u^*(T)B)\epsilon)C^*(T), \]
so
\[ \frac{dC}{d\epsilon}(T + \epsilon) = \exp((A + u^*(T)B)\epsilon)(A + u^*(T)B)C^*(T). \]

Let \( r = (w^*)' \frac{d}{d\epsilon} C(T + \epsilon) \big|_{\epsilon=0} v^*. \) Using the fact that \( T \) is locally maximal and arguing as in the proof of Theorem 4 yields \( r \leq 0 \), so
\[ 0 \geq (w^*)'((A + u^*(T)B)C^*(T))v^* \]
\[ = \rho^*(w^*)'(A + u^*(T)B)v^*. \] (5.14)

Now consider the control \( u : [0, T - \epsilon] \to [0, 1] \) defined by \( u(t) = u^*(t) \) for \( t \in [0, T - \epsilon] \). The corresponding solution at time \( T - \epsilon \) is just \( C^*(T - \epsilon) \), and its derivative is
\[ \frac{dC}{d\epsilon}(T - \epsilon) = -(A + u^*(T - \epsilon)B)C^*(T - \epsilon). \]

In particular,
\[ \frac{dC}{d\epsilon}(T - \epsilon) \big|_{\epsilon=0} = -(A + u^*(T)B)C^*(T). \]

Using the fact that \( T \) is locally maximal and arguing as in the proof of Theorem 4 yields
\[ (w^*)'(\frac{d}{d\epsilon} C(T - \epsilon) \big|_{\epsilon=0})v^* \leq 0 \]
so
\[ 0 \geq -(w^*)'((A + u^*(T)B)C^*(T))v^* \]
\[ = -\rho^*(w^*)'(A + u^*(T)B)v^*. \]
Combining this with (5.14) shows that \( H(p(T), q(T), u^*(T)) = 0 \) and this completes the proof.

**Example 4** Consider the PBCS in Example 2. A numerical search suggests that \( T = 0.1408484 \) is locally minimal (see Fig. 4.1), and
\[ C^*(T) = \exp((A + B)t_2)\exp(At_1), \]
with \( t_1 = 0.0649672, \ t_2 = 0.0758812 \). A calculation yields \( \rho(C^*(T)) = 0.95564055 \), and
\[ v^* = \begin{bmatrix} 0.58697487 & 0.70506070 & 0.39793203 \end{bmatrix}', \]
\[ w^* = \begin{bmatrix} 0.36454066 & 0.66099215 & 0.80411773 \end{bmatrix}'. \] (5.15)
Given (5.15), we numerically solved the differential equations in Thm. 4, and calculated \( m(t) \) and \( u^*(t) \). The switching function \( m(t) \) is shown in Fig. 5.2. Note
that the zeros of $m(t)$ on the interval $[0, T)$ are at $0, t_1$. A calculation shows that both $(w^*)'(Av^*)$ and $(w^*)'(Bv^*)$ are zero (the numerical values are of order $10^{-15}$), so
\[
H(0) = q'(0)(A + u^*(0)B)p(0) = \rho^*(w^*)'(A + u^*(0)B)v^* = 0.
\]

In the next section we describe several applications of Theorem 4.

6. Applications. Eq. (5.8) has an interesting application. To state it, we introduce two more definitions. We say that an optimal control $u^*$ is bang-bang if the set \( \{ t \in [0, T) : m(t) = 0 \} \) consists of a finite number of points $0 \leq s_1 < s_2 < \cdots < s_k < T$. We say that the bang-bang control is regular if $m(t)$ changes sign at $s_i, i = 1, \ldots, k$. In this case, we refer to each $s_i$ as a switching point. Clearly, a sufficient condition for a bang-bang control $u^*$ to be regular is that $\dot{m}(s_i) \neq 0, i = 1, \ldots, k$.

The latter condition plays an important role in deriving sufficient conditions for local optimality of bang-bang controls (see, e.g., [2; 29]).

**Corollary 1** Suppose that the conditions of Thm. 4 hold. Let $u^*$ be a regular bang-bang control with $k$ switching points $0 \leq s_1 < s_2 < \cdots < s_k < T$. Then $k$ is even.

**Proof.** Without loss of generality, assume that $s_1 \neq 0$, i.e. $m(0) \neq 0$. By (5.8) and the fact that $m(t)$ changes sign at each switching point, $k$ must be even.

Note that the results in Examples 3 and 4 agree with Corollary 1, as in these examples the number of switching points in $[0, T)$ is 4 and 2, respectively (see Figs. 5.1 and 5.2).

Motivated by the PASP, we now consider the PBCS (4.4) with $B = bc'$, $b, c \in \mathbb{R}^n$. This corresponds to a PLSS that switches between $\dot{x} = Ax$ and $\dot{x} = (A + bc')x$. 

![Fig. 5.2. Switching function $m(t)$ as a function of $t$.](image)
6.1. The case $B = bc'$. Recall that the switched system (1.1) is said to reduce if there exists a linear subspace $L$, with $L \neq \{0\}$ and $L \neq \mathbb{R}^n$, that is an invariant set for both $\dot{x} = A_0 x$ and $\dot{x} = A_1 x$. In our case, a standard assumption guaranteeing that the switched system is not reducible is the following (see, e.g. [9]).

**Assumption 2** The pair $(A, b)$ is controllable and the pair $(A, c)$ is observable.

**Proposition 1** Consider the PBCS (4.4) with $B = bc'$. Assume that Assumption 2 holds. If the conditions of Theorem 4 hold, then for any final time $T > 0$, any optimal control $u^* \in U$ is bang-bang.

The proof will use the following result.

**Lemma 1** [1, Ch. 15] Consider the ordinary differential equation

$$y^{(n)}(t) + \alpha_1(t)y^{(n-1)}(t) + \cdots + \alpha_n(t)y(t) = 0$$

defined on some time interval $t \in [\tau, \tau + \epsilon]$, with $\epsilon > 0$. Suppose that the $\alpha_i$s are measurable and bounded functions:

$$\beta_i = \max_{t \in [\tau, \tau + \epsilon]} |\alpha_i(t)| < \infty,$$

and that $\epsilon$ is sufficiently small so that

$$\sum_{j=1}^{n} \beta_j \epsilon^j < 1.$$

If $y$ is not identically zero, then it has no more than $n - 1$ zeros on $[\tau, \tau + \epsilon]$.

**Proof of Proposition 1.** Let $u^* \in U$ be an optimal control. The corresponding switching function is

$$m(t) = q'(t)Bp(t) = g(t)f(t),$$

where $f(t) = c'\bar{p}(t)$ and $g(t) = b'q(t)$. Suppose that $f(s) = 0$ for some time $s \in [0, T)$. Then

$$\dot{f}(s) = c'\bar{p}(s) = c'(A + u^*(s)bc')p(s) = c'Ap(s),$$

where the last equation follows from the fact that $f(s) = c'\bar{p}(s) = 0$. Inductively, $f^{(j)}(s) = c'A^jp(s)$. We will show that there exists $k \in \{0, 1, \ldots, n-2\}$ such that $f(s) = f^{(k)}(s) = \cdots = f^{(1)}(s) = 0$, but $f^{(k+1)}(s) \neq 0$, so $s$ is an isolated zero of $f$. Indeed, if this is not the case then all the vectors $c', c'A, \ldots, c'A^{n-1}$ are orthogonal to the vector $\bar{p}(s)$. But this is impossible since the pair $(A, c)$ is observable.

Similarly, the controllability assumption implies that $g(t)$ has isolated zeros. Hence, $m(t) = g(t)f(t)$ has isolated zeros. By Theorem 4, this implies that $[0, T) =$
$I_0 \cup I_1$, where $I_0$, $I_1$ are two disjoint unions of intervals with $u^*(t) = 0$ [$u^*(t) = 1$] for $t \in I_0$ [$t \in I_1$]. Then,

$$f^{(k)}(t) = \begin{cases} \epsilon'A^kp(t), & t \in I_0, \\ \epsilon'(A + bc')^kp(t), & t \in I_1. \end{cases}$$

Let $\det (sI - A) = s^n + \cdots + a_1s + a_0 \det(sI - (A + bc')) = s^n + \cdots + b_1s + b_0$ denote the characteristic polynomial of $A [A + bc']$. By Cayley-Hamilton,

$$f^{(n)}(t) + a_{n-1}f^{(n-1)}(t) + \cdots + a_0f(t) = 0, \quad t \in I_0.$$  

and

$$f^{(n)}(t) + b_{n-1}f^{(n-1)}(t) + \cdots + b_0f(t) = 0, \quad t \in I_1.$$  

This implies that $f$ satisfies condition (6.1) with $\beta_i = \max\{a_i, b_i\}$. By Lemma 1, $f$ has a finite number of zeros on any finite time interval. A similar argument shows that $g$ has a finite number of zeros on any finite time interval. This implies that $u^*$ is a bang-bang control.

Prop. 1 shows that under the stated conditions the number of switching points is always finite. The next result shows that if $u^*$ has a switching point, then it has at least four switching points. Surprisingly, perhaps, this holds for any $T > 0$, regardless of how small $T$ is.

**Theorem 6** Suppose that the conditions of Proposition 1 hold. Let $u^* : [0, T] \to [0, 1]$ be an optimal control. If $u^*$ is regular and includes a switching point, then it includes at least four switching points.

The proof requires several tools. The first is a linear coordinate transformation, introduced in [9]. To motivate this transformation, consider the SISO linear system

$$\begin{align*}
\dot{x} &= Ax + bu, \\
y &= c'x.
\end{align*}$$  

(6.2)

The transfer function of this system is

$$\frac{Y(s)}{U(s)} = c'(sI - A)^{-1}b,$$  

(6.3)

where $Z(s)$ denotes the Laplace transform of a function $z(t)$. The SISO system

$$\begin{align*}
\dot{\tilde{x}} &= A'\tilde{x} + cu, \\
y &= b'\tilde{x}.
\end{align*}$$  

(6.4)

has an identical transfer function. This suggests the following.

**Proposition 2** If the SISO system (6.2) is controllable and observable, then there exists an invertible matrix $P \in \mathbb{R}^{n \times n}$ such that

$$A = P^{-1}A'P, \quad c = Pb, \quad c' = b'P.$$  

(6.5)
This result was already used in the analysis of the ASP in [9]. For the sake of completeness, we include a proof of Proposition 2 in the Appendix.

We can now prove Theorem 6. Suppose that $u^*$ is a regular bang-bang control with two switching points $s_1 < s_2$ in $[0, T)$. Without loss of generality, assume that $s_1 = 0$, and that

$$m(t) < 0, \quad t \in (s_1, s_2). \quad (6.6)$$

Then for $t \in (s_1, s_2)$, $u^*(t) = 0$ and

$$m(t) = q'(t)bc'p(t)$$

$$= b' \exp(-A't)q(0)c' \exp(At)p(0)$$

$$= b' \exp(-A't)(\rho^* w^*)c' \exp(At)v^*$$

$$= \rho^* b'PP^{-1} \exp(-A't)PP^{-1}w^*c' \exp(At)v^*$$

$$= \rho^* c' \exp(-At)PP^{-1}w^*c' \exp(At)v^*, \quad (6.7)$$

where the last equation follows from (6.5).

To simplify this expression, we need to relate $PP^{-1}w^*$ to $v^*$. To do this, note that

$$q(T) = \exp(-(A + bc')s_{3,2}) \exp(-A's_{2,1})q(0),$$

where $s_{i,k} = s_i - s_k$ and $s_3 = T$. Hence,

$$q(0) = \exp(A's_{2,1}) \exp((A + bc')s_{3,2})q(T),$$

and combining this with (5.7) yields

$$\rho^* w^* = \exp(A's_{2,1}) \exp((A + bc')s_{3,2})w^*,$$

so

$$\rho^* PP^{-1}w^* = PP^{-1} \exp(A's_{2,1})PP^{-1} \exp((A + bc')s_{3,2})PP^{-1}w^*$$

$$= \exp(A's_{2,1}) \exp((A + bc')s_{3,2})PP^{-1}w^*.$$ \hspace{1cm} (6.8)

Multiplying both sides of this equation by $\exp((A + bc')s_{3,2})$ yields

$$\rho^* \exp((A + bc')s_{3,2})PP^{-1}w^* = C^*(T) \exp((A + bc')s_{3,2})PP^{-1}w^*.$$ 

Since the eigenvector $v^*$ of $C^*(T)$ corresponding to $\rho^*$ is unique (up to scaling), this implies that

$$\exp((A + bc')s_{3,2})PP^{-1}w^* = rv^*,$$ 

for some $r \in \mathbb{R} \setminus \{0\}$. Combining this with (6.8) yields

$$P^{-1}w^* = \frac{r}{\rho^*} \exp(A's_{2,1})v^*, \quad (6.9)$$

and substituting this in (6.7) gives

$$m(t) = rc' \exp(A(s_{2,1} - t))v^*c' \exp(At)v^*, \quad (6.10)$$

where the last equation follows from (6.5).
In particular,
\[ m(s_{2,1}/2) = m(s_2/2) = r \left( c' \exp(A(s_2/2))v^* \right)^2, \]
so (6.6) implies that \( r < 0 \).

Now consider a time \( t \in (s_2, s_3) \). Then
\[
\begin{align*}
m(t) &= b'q(t)c'\rho(t) \\
&= b' \exp(-(A + bc')(t - s_3))q(s_3)c' \exp((A + bc')(t - s_2))p(s_2) \\
&= b' \exp(-(A + bc')(t - s_3))w^c' \exp((A + bc')(t - s_2)) \exp(As_2)v^* \\
&= b'PP^{-1} \exp(-(A + bc')(t - s_3))PP^{-1}w^c' \exp((A + bc')(t - s_2)) \exp(As_2)v^*,
\end{align*}
\]
and using (6.5) and (6.9) yields
\[
m(t) = (r/\rho^*)c' \exp(-(A + bc')(t - s_3)) \exp(As_2)v^* c' \exp((A + bc')(t - s_2)) \exp(As_2)v^*,
\]
Let \( \tilde{t} = (s_3 + s_2)/2 \). Then
\[
m(\tilde{t}) = (r/\rho^*)c' \exp((A + bc')(s_3 - s_2)/2) \exp(As_2)v^* \bigg| \leq 0.
\]
This is a contradiction since \( s_2 \) is a switching point so \( m(t) > 0 \) for \( t \in (s_2, s_3) \). We conclude that a regular bang-bang control \( u^* \) cannot have two switching points. Hence, if \( u^* \) has a switching point, it must have at least three switching points. By Corollary 1, \( u^* \) has at least four switching points. This completes the proof of Theorem 6.

7. Stability Analysis. We now consider applications of Theorem 4 to stability analysis. From here on we assume that \( A + rB \) is Hurwitz for any \( r \in [0, 1] \). When \( A, B \) are Metzler, \( C(T, u) \) will be non-negative but not necessarily primitive, so we cannot immediately apply Theorem 4. The next result provides a simple trick to overcome this limitation. We use 1 to denote the \( n \times n \) matrix with all entries equal to one.

**Proposition 3** Consider the PBCS with matrices
\[
\bar{A} = A + \epsilon 1, \quad \bar{B} = B,
\]
with \( \epsilon > 0 \). Let \( \bar{C} \) denote the corresponding transition matrix. Then for any \( T > 0 \) and any \( u \in U \),
\[
\bar{C}(T, u) > 0, \quad \rho(\bar{C}(T, u)) \geq \rho(C(T, u)).
\]

**Proof.** Fix an arbitrary \( r \in [0, 1] \). The Metzler property implies that \( A + rB = pI + D \), with \( p = \min_i \{ a_{ii} + rb_{ii} \} \) and \( D \geq 0 \). Thus,
\[
\exp((\bar{A} + r\bar{B})t) = \exp((pI + D + \epsilon 1)t)
\]
\[
= \exp(pt) \sum_{i=0}^{\infty} \frac{(D + \epsilon 1)^i}{i!} t^i.
\]
Since \( D + \epsilon I \succ 0 \), we conclude that \( \exp((\bar{A} + r\bar{B})t) > 0 \) for any \( t > 0 \). The reachable set of a PBCS for piecewise constant controls is a dense subset of the reachable set for measurable controls, so this proves (7.2). To prove (7.3), note that

\[
\exp((\bar{A} + r\bar{B})t) = \exp(pt) \sum_{i=0}^{\infty} \frac{(D + \epsilon I)^i t^i}{i!} \\
\geq \exp(pt) \sum_{i=0}^{\infty} \frac{D^i t^i}{i!} \\
= \exp((A + rB)t),
\]

so \( \rho(\exp((\bar{A} + r\bar{B})t)) \geq \rho(\exp((A + rB)t)). \)

**Remark 5** Proposition 3 implies that the PBCS with the matrices in (7.1) always satisfies the conditions of Theorem 4 (as \( \bar{C}(T, u) > 0 \) implies that \( \bar{C}(T, u) \) is primitive), and that GAS of this PBCS implies GAS of the PBCS with the original matrices \( A, B \).

The next result demonstrates how the MP may be used to prove GAS of a PBCS (and thus GUAS of the corresponding PLSS).

**Corollary 2** Consider the PBCS (4.1). Suppose that \( A + rB \) is Hurwitz for any \( r \in [0, 1] \). If there exist \( \alpha, \beta \in \mathbb{R} \) such that

\[
\alpha A + \beta(A + B) > 0,
\]  
(7.4)

then the PBCS is GAS.

**Proof.** Consider the PBCS with matrices \( \bar{A}, \bar{B} \) in (7.1) with \( \epsilon > 0 \) sufficiently small so that \( \bar{A} + r\bar{B} \) is Hurwitz for any \( r \in [0, 1] \), and \( \alpha\bar{A} + \beta(\bar{A} + \bar{B}) > 0 \). Then \( \bar{C}^*(t) > 0 \) for any \( t > 0 \), so we can apply Theorem 5. Assume that some time \( \bar{T} > 0 \) is extremal. Since \( \bar{v}^*, \bar{w}^* > 0 \),

\[
(\bar{w}^*)'(\alpha\bar{A} + \beta(\bar{A} + \bar{B}))\bar{v}^* > 0.
\]

But since \( \bar{T} \) is extremal, \( (\bar{w}^*)'\bar{A}\bar{v}^* = (\bar{w}^*)'\bar{B}\bar{v}^* = 0 \). This contradiction shows that no time \( \bar{T} > 0 \) can be extremal.

Define \( \eta : \mathbb{R}_+ \to \mathbb{R}_+ \) by \( \eta(t) = \rho(\bar{C}^*(t)) \). It is not difficult to show that \( t \to \eta(t) \) is continuous. The fact that \( \bar{A} + r\bar{B} \) is Hurwitz for all \( r \in [0, 1] \) implies that there exists a sufficiently small \( \tau > 0 \) such that

\[
\eta(\tau) < 1.
\]  
(7.5)

Also, since \( \bar{C}(0) = I \),

\[
\eta(0) = 1.
\]  
(7.6)

As no time can be extremal, we conclude that \( \eta(T) \leq \eta(\tau) < 1 \) for any \( T \geq \tau \). Thm. 3 implies that the PBCS with matrices \( \bar{A}, \bar{B} \) is GAS. By Remark 5, this completes the proof.

It is important to note that condition (7.4) has a very simple interpretation. Since \( A, A + B \) are Hurwitz and Metzler, it is not possible that \( \alpha \) and \( \beta \) have the
same sign. Assume without loss of generality that $\alpha > 0$, and $\beta < 0$. Then (7.4) yields
\[
A > \frac{-\beta}{\alpha} (A + B)
\]
with $\frac{-\beta}{\alpha} > 0$, i.e. (up to a scaling that has no effect on stability)
\[
A > A + B.
\]
Since $A$ is Metzler and Hurwitz, there exists a diagonal matrix $D > 0$ such that $V(x) = x'Dx$ is a Lyapunov function for $\dot{x} = Ax$, and it straightforward to verify that (7.7) implies that $V$ is a common diagonal Lyapunov function (CDLF) for the PBCS (for more on CDLFs for PLSSs; see [46]).

7.1. The case $B = bc'$. We now consider the special case where $A + B$ is a rank one perturbation of $A$, i.e. $B = bc'$ for some $b, c \in \mathbb{R}^n$. Note that in this case finding a necessary and sufficient condition for GAS of the PBCS is equivalent to solving the PASP. Since a PBCS is GAS if and only if the associated linear switched system is GUAS, this is also equivalent to finding a necessary and sufficient condition for GUAS of (1.1) with $A_1, A_0$ Metzler, and $\text{rank}(A_1 - A_0) = 1$.

We say that a vector $y \in \mathbb{R}^n$ is sign-definite (SD) if either $y \geq 0$ or $y \leq 0$. It has been shown that if both $b$ and $c$ are SD, and $A + rbc'$ is Hurwitz and Metzler for any $r \in [0, 1]$, then $A$ and $A + bc'$ admit a common quadratic Lyapunov function and, therefore, the PBCS (4.1) with $B = bc'$ is GAS [3; 15; 40]. We prove a slightly stronger result.

**Proposition 4** Consider the PBCS (4.1) with $B = bc'$. Suppose that $A + rbc'$ is Hurwitz for any $r \in [0, 1]$, Assumption 2 holds, and that $c$ is SD or $b$ is SD. Then the PBCS is GAS.

**Proof.** For $k > 0$, let $C_k(t)$ denote the solution at time $t$ of the PBCS
\[
\dot{C} = (A + kbc'u)C, \\
C(0) = I.
\]
Seeking a contradiction, assume that (4.1) is not GAS. Then there exist $T > 0$ and $u^* \in \mathcal{U}$ such that $\rho(C_k(T)) \geq 1$. On the other-hand, for $k = 0$ the solution of (7.8) satisfies $\rho(C_k^0(T)) < 1$ (recall that $A$ is Hurwitz). By continuity, there exist $T > 0$ and $k \in (0, 1]$ such that $\rho(C_k^0(T)) = 1$. Then Theorem 4 yields
\[
p(T) = \rho^* p(0) = p(0),
\]
and
\[
c'A^{-1}\dot{p}(t) = c'A^{-1}(A + kbc'u(t))p(t) \\
= (1 + c'A^{-1}bku(t))c'p(t).
\]
Recall that $\det(A + bc'ku) = \det(A)(1 + c'A^{-1}bku)$.

\[\text{We may assume that } C_k^0(T) \text{ satisfies the condition in Theorem 4, as otherwise we can use the same idea as in Proposition 3.}\]
Since both $A$ and $A+kbc'$ are Metzler and Hurwitz, we conclude that $(1+c'A^{-1}bkv)>0$ for any $v \in [0,1]$. If $c \geq 0 \ [c \leq 0]$ then since $p(t)>0$ for $t \in [0,T]$, integrating (7.10) yields $c'A^{-1}(p(T)−p(0))>0$ $[c'A^{-1}(p(T)−p(0))<0]$. This contradicts (7.9). Thus, if $c$ is SD, then the PBCS is GAS. If $b$ is SD, a similar argument using the function $b'(A')^{-1}q$ shows that the PBCS is GAS.

We believe that Theorem 6 may also have interesting applications for stability analysis, but this issue is left for further research.

8. Conclusion. The transition matrix of a positive bilinear control system (PBCS) is non-negative (and non-singular), so its spectral radius is attained by a real and positive eigenvalue. Motivated by the problem of global asymptotic stability of a PBCS or, equivalently, the global uniform asymptotic stability of a PLSS, we considered the following optimal control problem. For a fixed final time, find a control that maximizes the spectral radius of the transition matrix. Our main result is a new first-order necessary condition for optimality in the form of a maximum principle (MP). The proof of this MP is based on combining the standard tool of a needle variation and the Perron-Frobenius theorem. We described several applications of this MP to stability analysis.

Possible directions for future research include: (1) developing a second-order MP for PBCSs; and (2) using the MP for analyzing other dynamical systems where non-negative matrices play an important role, for example, consensus problems (see, e.g. [50]).

Acknowledgements. We are grateful to Andrei Agrachev and Daniel Liberzon for helpful discussions. We thank Elena Valcher for thoroughly reading a draft of this paper and providing us with many helpful comments. We are grateful to the anonymous reviewers for their constructive comments.

Appendix. Proof of Proposition 2. Let $\det(sI−A)=s^n+an−1s^{n−1}+\cdots+a_0$ be the characteristic polynomial of $A$. By Cayley–Hamilton, $A^n+an−1A^{n−1}+\cdots+a_0I=0$. Since $(A,b)$ is controllable, the matrix

$$C=\begin{bmatrix} b & Ab & \ldots & A^{n−1}b \end{bmatrix}$$

is invertible. The transformation $z=C^{-1}x$ transforms (6.2) into the canonical form

$$\dot{z} = C^{-1}ACz + C^{-1}bu$$

$$=A_cz + e_1u,$$

where $A_c = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & -a_2 \\ \vdots \\ 0 & 0 & 0 & \ldots & 1 & -a_{n−1} \end{bmatrix}$, and $e_1 = [1 \ 0 \ \ldots \ 0]'$. Similarly, since $(A,c)$ is observable, the matrix

$$O = \begin{bmatrix} c & A'c & \ldots & (A')^{n−1}c \end{bmatrix}$$

is invertible. The transformation $r = O^{-1}\tilde{x}$ transforms (6.4) into the canonical form

$$\dot{r} = O^{-1}A'O + O^{-1}cu$$

$$=A_cr + e_1u.$$
Now it is straightforward to verify that all the equations in (6.5) hold for $P = OC^{-1}$.
For example,

\[
P^{-1}A'P = CA_cC^{-1} = A.
\]

References.


[40] ———, Stability analysis of positive linear switched systems: a variational ap-
L. FAINSHIL AND M. MARGALIOT

Proach, in Proc. 6th IFAC Symposium on Robust Control Design (ROCOND09), Haifa, Israel, 2009.


