Approximation of The Constrained Joint Spectral Radius via Algebraic Lifting

Xiangru Xu, Behcet Acikmese

Abstract

A linear constrained switching system is a discrete-time linear switched system whose switching sequences are constrained by a deterministic finite automaton. As a characterization of the asymptotic stability of a constrained switching system, the constrained joint spectral radius is difficult to compute or approximate. Using the semi-tensor product of matrices, we express dynamics of a deterministic finite automaton, an arbitrary switching system and a constrained switching system into their matrix forms, respectively, where the matrix expression of a constrained switching system can be seen as the matrix expression of a lifted arbitrary switching system. Inspired by this, we propose a lifting method for the constrained switching system, and prove that the constrained joint/generalized spectral radius of the constrained switching system is equivalent to the joint/generalized spectral radius of the lifted arbitrary switching system. Examples are provided to show the advantages of the proposed lifting method.

1 Introduction

Consider a finite set of matrices $\mathcal{A} = \{A_1, \ldots, A_m\}$ with $A_i \in \mathbb{R}^{n \times n}$, $i \in [m]$ where $[m] := \{1, 2, \ldots, m\}$. Dynamics of the discrete-time linear switched system associated with $\mathcal{A}$ are described as

$$ x_{k+1} = A_{\sigma_k} x_k $$

where $x_k \in \mathbb{R}^n$ is the state, and $\sigma_k \in [m]$ is the switching signal at time step $k$. As there is no constraint on the switching sequence, the system \cite{1} is called the arbitrary switching system and denote by $S(\mathcal{A})$ \cite{20}. We call the arbitrary switching system asymptotically stable if $\lim_{k \to \infty} x_k = 0$ for any $x_0 \in \mathbb{R}^n$ and any switching sequence $\sigma_0 \sigma_1 \ldots$ (see \cite{16, 17} and references therein for more details about linear switched systems).

The joint spectral radius (JSR) of $\mathcal{A}$ is defined as

$$ \rho(\mathcal{A}) = \limsup_{k \to \infty} \rho_k(\mathcal{A})^{1/k} $$

where

$$ \rho_k(\mathcal{A}) = \max_{\sigma \in [m]^k} \|A_\sigma\|, $$

---

*X. Xu and B. Acikmese are with the Department of Aeronautics & Astronautics, University of Washington, Seattle, WA, USA. Email: {xiangru, behcet}@uw.edu.*
\( \sigma = \sigma_0 \ldots \sigma_{k-1} \) is a switching sequence of length \( k \) with \( \sigma_0, \ldots, \sigma_{k-1} \in [m] \).

\( A_\sigma := A_{\sigma_{k-1}} \ldots A_{\sigma_0} \) is the product of \( k \) matrices, and \( \| \cdot \| \) is any given sub-multiplicative matrix norm on \( \mathbb{R}^{n \times n} \). The concept of JSR for a finite set of matrices is a natural generalization of the spectral radius for a single (square) matrix, and was first introduced by Rota and Strang in [21]. Because of the equivalence of the matrix norms in finite-dimensional vector spaces, the value of \( \rho(A) \) is independent of the choice of the matrix norm. The JSR found applications in many areas such as the continuity of wavelet functions, the capacity of codes, and the trackable graphs [12]. Particularly, the value of \( \rho(A) \) characterizes the asymptotic stability of the switched system (1), as (1) is asymptotically stable if and only if \( \rho(A) < 1 \). However, the value of \( \rho(A) \) is notoriously difficult to compute or approximate (see the NP-hardness and the undecidability results in [4, 22]). In the past decade, various methods for approximating \( \rho(A) \) have been proposed, such as using branch and bound [10], convex combination [3], lifted polytope [13], sum-of-squares [19, 15], and path-complete graph Lyapunov functions [1].

By replacing the norm in (3) with the spectral radius, Daubechies and Lagarias introduced the concept of generalized spectral radius (GSR) of \( A \) in [9]. Specifically, the GSR of \( A \) is defined as

\[
\bar{\rho}(A) = \limsup_{k \to \infty} \bar{\rho}_k(A)^{1/k}
\]

where

\[
\bar{\rho}_k(A) = \max_{\sigma \in [m]^k} \rho(A_\sigma).
\]

The Berger-Wang Theorem proves that the JSR and the GSR of \( A \) are actually equivalent, i.e., \( \rho(A) = \bar{\rho}(A) \) [2].

The switching sequence \( \sigma \) of the switched system (1) can be subject to certain constraints. For instance, the switching signal needs to satisfy the Markovian-like property (i.e., \( \sigma_k \) that is allowable is dependent on \( \sigma_{k-1} \) [13] [6]), or it needs to be accepted by a Muller automaton [24]. Following [20], in this paper, we consider the switching sequences that are constrained by a deterministic finite automaton (DFA).

**Definition 1.** A deterministic finite automaton \( M \) is a 3-tuple \( (Q, U, f) \) where \( Q \) is a finite set of states, \( U \) is a finite set of input symbols, \( f : X \times U \to X \) is the transition function.

Note that in the definition above, the initial state and the final states of \( M \) are not specified, and the transition function is a partial function that may not be defined for all state-input pairs. Without loss of generality, we assume that \( M \) is alive (i.e., for each state \( q \in Q \) there is at least one \( u \in U \) such that \( f(q, u) \) is defined). A finite input sequence \( u_1 u_2 \ldots u_k \) is said to be accepted by \( M \) if there exists a finite state sequence \( q_1 q_2 \ldots q_k \) such that \( f(q_i, u_i) \) is defined for all \( i \in [k] \); an infinite input sequence accepted by \( M \) is defined similarly by taking \( k = \infty \). The set of input sequences accepted by \( M \) forms a symbolic dynamical system that is called sofic shift (see Section 1.5 of [18]), and is denoted by \( L(M) \).

A DFA \( M \) can be considered as a directed and labeled graph \( M(V, E) \) where \( V \) is the set of nodes and \( E \) is the set of edges. The edge \((v, w, s) \in E \) if there is an edge from the node \( v \) to \( w \) via the label \( s \). The constrained switching system,
denoted as \( S(\mathcal{A}, \mathcal{M}) \), is the linear switched system\(^1\) where \( A_i \in \mathcal{A} \) for \( i \in [m] \), the switching sequence \( \sigma = \sigma_0 \sigma_1 \sigma_2 \ldots \) satisfies \( \sigma \in \mathcal{L}(\mathcal{M}) \), and \( \sigma_k \in [m] \).

The concept of JSR can be naturally generalized to the case when the switching sequences are constrained by a DFA. Specifically, the constrained joint spectral radius (CJSR) of \( S(\mathcal{A}, \mathcal{M}) \) is defined as

\[
\rho(\mathcal{A}, \mathcal{M}) = \limsup_{k \to \infty} \rho_k(\mathcal{A}, \mathcal{M})^{1/k}
\]

where

\[
\rho_k(\mathcal{A}, \mathcal{M}) = \max_{\sigma \in [m]^k, \sigma \in \mathcal{L}(\mathcal{M})} \| A_{\sigma} \|.
\]

Similarly, the constrained generalized spectral radius (CGSR) of \( S(\mathcal{A}, \mathcal{M}) \) is defined as

\[
\bar{\rho}(\mathcal{A}, \mathcal{M}) = \limsup_{k \to \infty} \bar{\rho}_k(\mathcal{A}, \mathcal{M})^{1/k}
\]

where

\[
\bar{\rho}_k(\mathcal{A}, \mathcal{M}) = \max_{\sigma \in [m]^k, \sigma \in \mathcal{L}(\mathcal{M})} \rho(\mathcal{A}_{\sigma}).
\]

The value of \( \rho(\mathcal{A}, \mathcal{M}) \) is independent of the choice of the matrix norm in \( \rho_k(\mathcal{A}, \mathcal{M}) \), and it characterizes the asymptotic stability of the constrained switching system \( S(\mathcal{A}, \mathcal{M}) \) as \( S(\mathcal{A}, \mathcal{M}) \) is asymptotically stable if and only if \( \rho(\mathcal{A}, \mathcal{M}) < 1 \)\(^{20}\).

Due to the constraint on the switching sequences, the computation or approximation of \( \rho(\mathcal{A}, \mathcal{M}) \) is more difficult than \( \rho(\mathcal{A}) \), with only a few results known in the literature: in \( \cite{20} \), the problem of approximating \( \rho(\mathcal{A}, \mathcal{M}) \) was reduced to finding a good multinorm (i.e., a set of norms that are defined for each node of \( \mathcal{M}(V,E) \)), where an arbitrarily accurate approximation can be obtained by solving a semi-definite program and using the quadratic-type multinorm; in \( \cite{15} \), an algorithm that generates a sequence of matrices with asymptotic growth rate close to the CJSR was proposed, based on the dual solution of a sum-of-squares optimization program. The linear switched system whose switching sequences are constrained by a Muller automaton was considered in \( \cite{24} \), where a lifting method based on the Kronecker product was proposed and used to show how different notions of stability are related; the linear switched system whose switching sequences are constrained by a given square matrix was considered in \( \cite{8} \) and \( \cite{14} \), where the Markovian joint spectral radius was discussed and the Markovian analog of the Berger-Wang formula was derived.

In this paper, we propose a lifting method that transforms a constrained switching system \( S(\mathcal{A}, \mathcal{M}) \) equivalently into an arbitrary switching system \( S(\mathcal{A}_M) \), such that the approximation of the CJSR/CGSR of \( S(\mathcal{A}, \mathcal{M}) \) can be converted into the approximation of the JSR/GSR of \( S(\mathcal{A}_M) \), where many off-the-shelf algorithms can be leveraged (see Figure 1 for an illustration of the main results).

The contributions of the paper are summarized as follows: 1) we propose a uniform matrix expression for the arbitrary switching system and the constrained switching system by using the semi-tensor product (STP) of matrices; 2) we prove a version of the Berger-Wang formula for the constrained switching system, which shows that \( \rho(\mathcal{A}, \mathcal{M}) = \bar{\rho}(\mathcal{A}, \mathcal{M}) = \rho(\mathcal{A}_M) = \bar{\rho}(\mathcal{A}_M) \); 3) we discuss
the connection of the STP-based method with some other lifting methods in
the approximation of CJSR of the constrained switching system. The remainder
of the paper is organized as follows. In Section 2, we introduce some prelimi-
naries about STP. In Section 3, we present the STP-based matrix formulation for
the arbitrary switching system, the DFA and the constrained switching system.
In Section 4, we prove that the CJSR/GCSR of $S(A, M)$ and the JSR/GSR of
$S(A_M)$ are all equivalent. In Section 5, we discuss the connection of
the STP-based method with some other lifting methods. Finally, some concluding
remarks are given in Section 6.

Given matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$, their conventional matrix product $AB$ requires $n = p$. The Kronecker product of $A$ and $B$, however, has no such
dimensional restriction on $n$ and $p$.

**Definition 2.** Let $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$, their
Kronecker product is defined as

$$A \otimes B := \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}. $$

The following two properties of the Kronecker product will be used in later
sections:

- Given matrices $A \in \mathbb{R}^{m_A \times n_A}$, $B \in \mathbb{R}^{m_B \times n_B}$, $C \in \mathbb{R}^{n_A \times n_C}$, $D \in \mathbb{R}^{n_B \times n_D}$,
  it holds that

  $$(A \otimes B)(C \otimes D) = (AC) \otimes (BD).$$

- Given two matrices $A \in \mathbb{R}^{p \times p}$ and $B \in \mathbb{R}^{q \times q}$, if $\lambda_1, \ldots, \lambda_p$ are the eigen-
  values of $A$ and $\mu_1, \ldots, \mu_q$ are the eigenvalues of $B$, then the eigenvalues
  of $A \otimes B$ are $\lambda_i \mu_j$ for $i = 1, \ldots, p$ and $j = 1, \ldots, q.$
Similar to the Kronecker product, the semi-tensor product of matrices can be also defined for two matrices with arbitrary dimensions $[5, 6, 7]$.

**Definition 3.** [7] Given two matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$, their semi-tensor product is defined as

$$A \Join B := (A \otimes I_{s/n}) (B \otimes I_{s/p})$$

where $s$ is the least common multiple of $n$ and $p$, and $\otimes$ is the Kronecker product.

Clearly, STP becomes the conventional matrix product when $n = p$, and STP becomes the Kronecker product when $n$ and $p$ are co-prime. Moreover, STP not only has the properties of associativity and distributivity as the conventional matrix product, but also has some unique properties as shown below (see [7] for more details):

- Given a column vector $x \in \mathbb{R}^n$ and a matrix $A$, it holds that
  $$x \Join A = (I_n \otimes A) \Join x.$$ (10)

- Given a column vector $x \in \Delta_n$, there exists a matrix $\Phi_n = \text{diag}(\delta^1_n, \delta^2_n, \ldots, \delta^n_n) \in \mathcal{L}_{n^2 \times n}$ such that
  $$x \Join x = \Phi_n x.$$ (11)

- Given two column vectors $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, there is a matrix $W_{[n,m]} = [\delta^1_m \Join \delta^1_n, \ldots, \delta^m_m \Join \delta^1_n, \ldots, \delta^1_m \Join \delta^m_n, \ldots, \delta^m_m \Join \delta^m_n] \in \mathcal{L}_{mn \times mn}$ such that
  $$W_{[n,m]} \Join x \Join y = y \Join x.$$ (12)

### 3 Semi-Tensor Product Formulation of the Arbitrary and Constrained Switching Systems

#### 3.1 STP Formulation of the Arbitrary Switching System

In this subsection, we present a STP-based matrix formulation for the arbitrary switching system [1].

Given a finite set of matrices $A = \{A_1, \ldots, A_m\}$ where $A_i \in \mathbb{R}^{n \times n}$, $i \in [m]$, we define a matrix $H$ as

$$H = [A_1, \ldots, A_m] \in \mathbb{R}^{n \times nm}. \quad (13)$$

For any $i \in [m]$, we identify $i$ with $\delta^i_m$, i.e., $i \sim \delta^i_m$. For any $\sigma_k$, which is the switching sequence of [1] at time step $k$, we define the vector form of $\sigma_k$ as a column vector $\sigma(k) \in \Delta_m$ where $\sigma(k) = \delta^i_m$ when $\sigma_k = i$, with $i \in [m]$ an arbitrary number. In other words, we identify $\sigma_k = i \in [m]$ with its vector form $\sigma(k) = \delta^i_m \in \Delta_m$, denoted as $\sigma_k \sim \sigma(k)$. We define $x(k) \in \mathbb{R}^n$ as the state of the switched system [1] by letting $x(k) = x_k$. In the rest of the paper, we will use $\sigma_k$ and $\sigma(k)$, $x_k$ and $x(k)$ interchangeably when there is no confusion.

Under these new notations, dynamical equation [1] for the arbitrary switching system $S(A)$ can be written into its equivalent matrix expression with the help of STP.
Proposition 1. Dynamics of $S(A)$ can be written equivalent as

$$x(k+1) = H \times \sigma(k) \times x(k)$$

(14)

where $H$ is given in (13), $x(k) \in \mathbb{R}^n$ is the state, and $\sigma(k) \in \Delta_m$ is the vector form of the input.

Proof. By the definition of STP, for any $k \in \mathbb{Z}_{\geq 0}$, it holds that $H \times \sigma(k) = A_{\sigma_k}$ where $\sigma(k) \sim \sigma_k$, $\sigma_k \in [m]$. Hence, (14) is equivalent to $x(k+1) = A_{\sigma_k} x(k)$, which is exactly the dynamics of system (1).

A finite switching sequence $\sigma = \sigma_0 \ldots \sigma_{k-1} \in [m]^k$ can be expressed equivalently into its vector form $\hat{\sigma} = \kappa_{\sigma_0} \ldots \kappa_{\sigma_{k-1}} \in \Delta_{m^k}$ where $i \in \{0, 1,\ldots, k-1\}$. Specifically, supposing that $\sigma_i = j_i$ where $j_i \in [m]$, $i \in \{0, 1,\ldots, k-1\}$, then $\sigma_i \sim \delta_{i}^{j_i}$ and the sequence $\sigma = \sigma_0 \ldots \sigma_{k-1}$ is identified with its vector form $\hat{\sigma} := \delta_{m^k}^{\sigma_0} \times \ldots \times \delta_{m^k}^{\sigma_{k-1}} \in \Delta_{m^k}$ where

$$\tau = 1 + \Sigma_{i=1}^{k} (j_{k-i} - 1)m^{k-i} \in [m^k].$$

(15)

Conversely, given a vector $\hat{\sigma} := \delta_{m^k}^{\sigma_0} \in \Delta_{m^k}$ where $\sigma \in [m]^k$, a set of numbers $j_0, \ldots, j_{k-1} \in [m]$ satisfying (15) can be uniquely determined, which corresponds to a switching sequence $\sigma = \sigma_0 \ldots \sigma_{k-1} \in [m]^k$. Hence, there is a one-to-one correspondence between a finite switching sequence $\sigma$ and its vector form $\hat{\sigma}$.

For any $k \geq 2$, from (14) and property (10), we have

$$x(k) = H \times \sigma(k-1) \times H \times \sigma(k-2) \times \ldots \times H \times \sigma(0) \times x(0)$$

$$= \hat{H}_k \times_{i=0}^{k-1} H \times \sigma(0) \times x(0)$$

(16)

where

$$\hat{H}_k = H \times_{i=1}^{k-1} (I_{m^i} \otimes H).$$

(17)

Noting that the matrix $\hat{H}_k$ in (17) has sizes $n \times nm^k$, we can partition it into $m^k$ sub-matrices as follows:

$$\hat{H}_k = [\hat{H}_{k_1}, \hat{H}_{k_2}, \ldots, \hat{H}_{km^k}]$$

(18)

where $\hat{H}_{ki} \in \mathbb{R}^{n \times n}$, $i \in [m^k]$. Given an arbitrary finite switching sequence $\sigma = \sigma_0 \ldots \sigma_{k-1} \in [m]^k$, from (10) we have $x(k) = A_{\sigma_k} x(0)$ where $A_{\sigma} = A_{\sigma_{k-1}} \ldots A_{\sigma_0} \in \mathbb{R}^{n \times n}$. If $\sigma_i = j_i$ where $j_i \in [m], i \in \{0, 1,\ldots, k-1\}$, then $\hat{H}_{ki} = A_{\sigma_i}$ where $\tau$ is given by (15). Therefore, $\hat{H}_k$ consists of $A_{\sigma_i}$ for all possible switching sequences $\sigma$.

Now let us consider computing the JSR/GSR of $S(A)$ using its matrix expression. It was shown in (12) that the following inequality holds for any $k \in \mathbb{Z}_{\geq 0}$:

$$\bar{\rho}_{k}(A)^{1/k} \leq \rho(A) = \rho(A) \leq \rho_{k}(A)^{1/k}.$$  

(19)

Based on (19), the value of $\rho(A)$ can be approximated to an arbitrary accuracy by increasing the value of $k$. It is not hard to see that $\rho_k(A) = \max_{i \in [m^k]} \| \hat{H}_{ki} \|$ and $\rho_k(A) = \max_{i \in [m^k]} \rho(\hat{H}_{ki})$. Hence, the inequality (19) results in the following proposition.
Proposition 2. Given a finite set of matrices $\mathcal{A} = \{A_1, \ldots, A_m\}$ where $A_i \in \mathbb{R}^{n \times n}$, $i \in [m]$, it holds that

$$\max_{i \in [m^k]} \rho(H_{ki})^{1/k} \leq \rho(\mathcal{A}) \leq \max_{i \in [m^k]} \|H_{ki}\|^{1/k} \tag{20}$$

where $H_{ki}$ is given in (18) with $H_k$ given in (17), and $\|\cdot\|$ is any given submultiplicative matrix norm.

The sizes of the matrix $H_k$ grow exponentially with $k$, which makes it difficult to compute $H_k$ and approximate $\rho(\mathcal{A})$ when $k$ is large.

Remark 1. An expression for $x(k)$ that is equivalent to (16) can be derived for any $k \geq 2$ as follows:

$$x(k) = HW_{[n, m]} \times x(k-1) \times \sigma(k-1)$$

$$= (HW_{[n, m]})^k \times x(0)^{\sigma(i)} \times k^{i=0} \sigma(i)$$

$$= H_k' \times x(0) \tag{21}$$

where $H_k' = (HW_{[n, m]})^k W_{[m, n]}$ and the swap matrix in (12) are used. Note that the switching sequences $\sigma(i)$ in (16) and (21) are multiplied in different orders, which means that they have different vector forms. The matrix $H_k'$ in (21) is different from the matrix $H_k$ in (16), but it also consist of $\mathcal{A}$ for all possible switching sequences $\sigma$, and therefore, can be used to compute $\rho(\mathcal{A})$ as in (20). Moreover, the matrix $(HW_{[n, m]})^k \times x(0)$ consists of the set of states that are reachable from the initial state $x(0)$ for all possible switching sequences of length $k$.

Example 1. Consider the finite set of matrices $\mathcal{A}$ given in [27], which describes a linear switched system that may experience controller failures. This set $\mathcal{A}$ is given by $\mathcal{A} = \{A_1, A_2, A_3, A_4\}$, where

$$A_1 = \begin{pmatrix} 0.94 & 0.56 \\ -0.35 & 0.73 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0.94 & 0.56 \\ 0.14 & 0.73 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 0.94 & 0.56 \\ -0.35 & 0.46 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0.94 & 0.56 \\ 0.14 & 0.46 \end{pmatrix}.$$

The matrix form of the dynamics of the arbitrary switching system $S(\mathcal{A})$ is $x(k+1) = H \times \sigma(k) \times x(k)$, where $x(k) \in \mathbb{R}^2$, $\sigma(k) \in [4]$, and $H = [A_1, A_2, A_3, A_4]$. Since $\rho(A_2) \approx 1.1340$ and $\rho(A_4) \approx 1.0688$, it is clear that $\rho(\mathcal{A}) > 1$ and the system $S(\mathcal{A})$ is unstable under arbitrary switching. By using (20) and letting $k = 7$, it can be calculated\(^1\) that $1.1340 \leq \rho(\mathcal{A}) \leq 1.1667$.

3.2 STP Formulation of the DFA

In this subsection, we revisit the STP-based matrix expression for the DFA that was proposed in [25, 26, 27].

Consider a DFA $\mathcal{M} = (Q, U, f)$ where $|Q| = \ell, |U| = m$. Without loss of generality, we assume that $Q = \{q_1, \ldots, q_\ell\}, U = \{1, \ldots, m\}$. Identify the state \(^1\)The calculation can be conducted using the STP toolbox for Matlab, which is available in [http://lsc.amss.ac.cn/~dcheng/](http://lsc.amss.ac.cn/~dcheng/)
$q_i \in Q$ with its vector form $\delta_i^f$ where $i \in [\ell]$ (denoted as $q_i \sim \delta_i^f$), and identify the input $j \in U$ with its vector form $\delta_m^j$ where $j \in [m]$ (denoted as $j \sim \delta_m^j$).

Then $X$ and $U$ are identified with the sets $\Delta_\ell$ and $\Delta_m$, respectively.

Define the transition structure matrix $F_j \in \mathbb{R}^{\ell \times \ell}$ associated with the input $j$ as

$$F_j(s,t) = \begin{cases} 1, & \text{if } \delta_t^s = f(\delta_t^j, \delta_m^j); \\ 0, & \text{otherwise.} \end{cases}$$

(22)

Define the transition structure matrix of $M$ as

$$F = [F_1, F_2, \ldots, F_m] \in \mathbb{R}^{\ell \times n\ell}.$$  

(23)

The DFA $M$ can be seen as a discrete-time dynamical system. Specifically, given an initial state $q_{i_0}$ and an input sequence $\sigma = \sigma_0 \sigma_1 \cdots$, $M$ evolves according to $q_{j_{i+1}} = f(q_{j_i}, \sigma_i)$ if the transition function $f(q_{j_i}, \sigma_i)$ is defined, where $j_0, j_1, \cdots \in [\ell], i_0, i_1, i_2, \cdots \in [m]$. If $f(q_{j_i}, \sigma_i)$ is not defined for some $i$, then the state $q_{j_{i+1}}$ is undefined and therefore, we let $q_{j_i} \equiv \delta_i^j$ for any $s \geq t + 1$.

Define $\sigma(k) \in \Delta_m$ and $q(k) \in \Delta_\ell^k$ as the vector forms of the state and the input of $M$ at step $k$, respectively, where $\sigma(k) = \delta_m^k$ for some $\kappa \in [m]$ if the input is $\delta_m^k$ at step $k$, $q(k) = \delta_\ell^k$ for some $s \in [\ell]$ if the state is $\delta_\ell^k$ at step $k$ and $q(k) = \delta_\ell^j$ if the state is undefined.

**Proposition 3.** [20, 21] The matrix form of the dynamics of $M$ is

$$q(k + 1) = F \times \sigma(k) \times q(k)$$

where $F$ is given in (23), $q(k) \in \Delta_\ell^k$ and $\sigma(k) \in \Delta_m$ are the vector forms of the state and input of $M$, respectively.

Similar to (16), for any $k \geq 2$ we have $q(k) = \tilde{F}_k \times \sigma(k - 1) \times q(0)$ where $\tilde{F}_k = F \times i_0 = (I_m \otimes F)$. Partition the matrix $\tilde{F}_k$ into $m^k$ sub-matrices as $\tilde{F}_k = [\tilde{F}_{k1}, \tilde{F}_{k2}, \ldots, \tilde{F}_{km}]$ where $\tilde{F}_{ki} \in \mathbb{R}^{n \times n}, i \in [m^k]$. Given an arbitrary finite switching sequence $\sigma = \sigma_0 \cdots \sigma_{k-1} \in [m]^k$, from (24) we have $q(k) = F_{\sigma} q(0)$ where $F_{\sigma} = F_{\sigma_{k-1}} \cdots F_{\sigma_0} \in \mathbb{R}^{n \times n}$. If the vector form of $\sigma$ is $\delta_m^{s_\ell}$ where $s \in [m^k]$, then it holds that $F_{\sigma} = \tilde{F}_{ks}$. Then the following corollary follows.

**Corollary 1.** Given a switching sequence $\sigma = \sigma_0 \cdots \sigma_{k-1} \in [m]^k$, $\sigma \in L(M)$ if and only if $F_{\sigma_{k-1}} \cdots F_{\sigma_0} \neq 0$.

**Example 2.** Consider the DFA $M = (Q, U, f)$ in [20], where $Q = \{q_1, q_2, q_3, q_4\}$, $U = \{1, 2, 3, 4\}$, and its transitions are shown in Figure 2. This DFA is designed to describe the constraint of the switching sequence, which implies that the same controller for the plant can not fail twice in a row (see [20] for more details). The matrix form of the dynamics of $M$ is $q(k+1) = F \times \sigma(k) \times q(k)$, where $q(k) \in \Delta_\ell^k$, $\sigma(k) \in [4]$, and the transition structure matrix of $M$ is $F = [F_1, F_2, F_3, F_4]$ with $F_1 = \delta_3[3, 3, 3, 3], F_2 = \delta_3[0, 1, 1, 0], F_3 = \delta_2[2, 0, 2, 0], F_4 = \delta_4[0, 0, 4, 0]$. Given an input sequence $\sigma = 231$, its vector form is $\delta_3 \times \delta_3 \times \delta_4 = \delta_4[0, 3, 3, 0]$. Calculate the matrix $\delta_4$, from which the 10-th block submatrix is $\delta_4[0, 3, 3, 0]$. Then this submatrix $\delta_4[0, 3, 3, 0]$ can be interpreted as follows: with the input sequence $\sigma = 231$, $M$ transitions to $q_3$ if it starts from $q_2$ or $q_3$, and the transition is not defined if it starts from $q_1$ or $q_4$. This can be easily verified by Figure 3.
3.3 STP Formulation of the Constrained Switching System

In this subsection, we express dynamics of a constrained switching system $S(A, \mathcal{M})$ into its matrix form.

Consider a finite set of matrices $A = \{A_1, \ldots, A_m\}$ where $A_i \in \mathbb{R}^{n \times n}$, $i \in [m]$, and a DFA $\mathcal{M} = (Q, U, f)$ where $|Q| = \ell$, $|U| = m$. Recall that dynamics of $S(A)$ and $\mathcal{M}$ are expressed into their matrix forms (14) and (24), respectively, which are restated below:

$$
\begin{align*}
x(k + 1) &= H \ast \sigma(k) \ast x(k), \\
q(k + 1) &= F \ast \sigma(k) \ast q(k).
\end{align*}
$$

Then let $\xi(k)$ be the state of $S(A, \mathcal{M})$ at time step $k$, which is defined as

$$\xi(k) = q(k) \otimes x(k) \in \mathbb{R}^{n\ell}. \quad (25)$$

Recalling the definition of STP, (25) is equal to

$$\xi(k) = q(k) \otimes x(k).$$

Clearly, $\xi \in \mathbb{R}^{n\ell}$ is a lifting of the state $x \in \mathbb{R}^n$ and has a block structure: if $q(k) = \delta_s^s$ for some $s \in [\ell]$, then the $s$-th block of $\xi(k)$ is equal to $x(k)$ with all the other blocks equal to zeros; if $q(k) = \delta_0^0$, then $\xi(k)$ is equal to $\delta_{n\ell}^0$. Then we have the following theorem.

**Theorem 1.** The matrix form of the dynamics of $S(A, \mathcal{M})$ is

$$\xi(k + 1) = \Phi \ast \sigma(k) \ast \xi(k) \quad (26)$$

where $\sigma(k) \in \Delta_m$ is the vector form of the input, $\xi(k)$ is the vector form of the state defined in (25), and $\Phi$ is the transition structure matrix that is defined as

$$\Phi = [\Phi_1, \ldots, \Phi_m]$$

with

$$\Phi_i = F_i \otimes A_i, \ \forall i \in [m], \quad (27)$$

and $F_i$ given in (23).
Proof. Supposing that \( \sigma(k) = \kappa \) where \( \kappa \) is an arbitrary number satisfying \( \kappa \in [m] \), it holds that \( q(k + 1) = F_\kappa q(k) \) and \( x(k + 1) = A_\kappa x(k) \). Therefore,

\[
\xi(k + 1) = F_\kappa q(k) \times A_\kappa x(k) \\
= F_\kappa \times (I_m \otimes A_\kappa) \times q(k) \times x(k) \\
= (F_\kappa \otimes I_n)(I_m \otimes A_\kappa)\xi(k) \\
= (F_\kappa \otimes A_\kappa)\xi(k) \\
= \Phi_\kappa \xi(k)
\]

where the second equality uses property \( \text{(10)} \), the third equality uses the definition of STP, and the fourth equality uses the mix-product property of the Kronecker product. Noting that \( \Phi \times \sigma(k) = \Phi_\kappa \), the conclusion follows immediately. \( \square \)

Define a finite set \( \mathcal{A}_M \subset \mathbb{R}^{n \times n} \) that consists of matrices \( \Phi_i \) given in \( \text{(27)} \), i.e.,

\[
\mathcal{A}_M = \{ \Phi_1, \ldots, \Phi_m \}
\]

Associated with \( \mathcal{A}_M \) is an arbitrary switching system \( S(\mathcal{A}_M) \) whose dynamics are given by \( \xi(k + 1) = \Phi_\sigma k \xi(k) \) where \( \xi_k \in \mathbb{R}^{n} \), \( \sigma_k \in [m] \), and \( \Phi_i \in \mathcal{A}_M \) for \( i \in [m] \).

We point out that \( S(\mathcal{A}_M) \) and \( S(\mathcal{A}, M) \) have the same transition structure matrix in their matrix expressions, and the arbitrary switching system \( S(\mathcal{A}_M) \) can be considered as a lifted system of the constrained switching system \( S(\mathcal{A}, M) \). In the next section, we will show that this lifting method is rather useful, and the CJSR/CGSR of \( S(\mathcal{A}, M) \) and the JSR/GSR of \( S(\mathcal{A}_M) \) are actually equivalent.

4 Equivalence of the CJSR/CGSR of \( S(\mathcal{A}, M) \) and the JSR/GSR of \( S(\mathcal{A}_M) \)

The following theorem is the main result of this paper, which is a version of the Berger-Wang formula for the constrained switching system \( S(\mathcal{A}, M) \). It shows that the CJSR and CGSR of \( S(\mathcal{A}, M) \), the JSR and GSR of \( S(\mathcal{A}_M) \) are all equivalent.

Theorem 2. The following equality holds:

\[
\rho(\mathcal{A}, M) = \bar{\rho}(\mathcal{A}, M) = \rho(\mathcal{A}_M) = \bar{\rho}(\mathcal{A}_M)
\]

Proof. For a matrix \( S \in \mathbb{R}^{n \times n} \), consider \( S \) as a \( \ell \times \ell \) block matrix \( S = (s_{ij}) \) with each block \( s_{ij} \in \mathbb{R}^{n \times n} \). Define the following function \( \| \cdot \| : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}_{\geq 0} \) (inspired by the function given in \( \text{(11)} \)):

\[
\| S \| = \max_{j \in [\ell]} \sum_{i=1}^{\ell} \| s_{ij} \|
\]

where \( \| \cdot \| \) is any given sub-multiplicative norm defined on \( \mathbb{R}^{n \times n} \). We claim that the function \( \| \cdot \| \) is a sub-multiplicative norm on \( \mathbb{R}^{n \times n} \). Indeed, \( \| \cdot \| \) is
Recalling (3) and (7), \( \rho \) is sub-additive can be easily seen from the fact that \( \|a_{ij} + b_{ij}\| \leq \|a_{ij}\| + \|b_{ij}\| \) for any \( i, j \in [\ell] \), and \( \|\cdot\| \) is sub-multiplicative because

\[
\|AB\| = \max_{j \in [\ell]} \sum_{i=1}^{\ell} \|\sum_{k=1}^{\ell} a_{ik}b_{kj}\|
\]

\[
\leq \max_{j \in [\ell]} \sum_{k=1}^{\ell} \sum_{i=1}^{\ell} \|a_{ik}\| \|b_{kj}\|
\]

\[
= \max_{j \in [\ell]} \left( \sum_{k=1}^{\ell} \|a_{ik}\| \right) \|b_{kj}\|
\]

\[
\leq \max_{j \in [\ell]} \left( \sum_{k=1}^{\ell} \|A\| \|b_{kj}\| \right)
\]

\[
= \|A\| \left( \max_{j \in [\ell]} \sum_{k=1}^{\ell} \|b_{kj}\| \right)
\]

\[
= \|A\| \|B\|
\]

Recalling (3) and (7), \( \rho_k(\mathcal{A}_M) \) is defined as \( \rho_k(\mathcal{A}_M) = \max_{\sigma \in [m]^k} \|\Phi_{\sigma}\| \) and \( \bar{\rho}_k(\mathcal{A}, \mathcal{M}) \) is defined as \( \bar{\rho}_k(\mathcal{A}, \mathcal{M}) = \max_{\sigma \in [m]^k, \sigma \in L(\mathcal{M})} \|A_{\sigma}\| \), where \( \Phi_{\sigma} = \Phi_{\sigma_{k-1}} \ldots \Phi_{\sigma_0}, A_{\sigma} = A_{\sigma_{k-1}} \ldots A_{\sigma_0} \) and \( \sigma = \sigma_0 \ldots \sigma_{k-1} \) is any switching sequence of length \( k \). From (27), we have

\[
\Phi_{\sigma} = (F_{\sigma_{k-1}} \odot A_{\sigma_{k-1}})(F_{\sigma_{k-2}} \odot A_{\sigma_{k-2}}) \ldots (F_{\sigma_0} \odot A_{\sigma_0})
\]

\[
= (F_{\sigma_{k-1}} F_{\sigma_{k-2}} \ldots F_{\sigma_0}) \odot (A_{\sigma_{k-1}} A_{\sigma_{k-2}} \ldots A_{\sigma_0})
\]

\[
= F_{\sigma} \odot A_{\sigma}
\]

(31)

where \( F_{\sigma} := F_{\sigma_{k-1}} F_{\sigma_{k-2}} \ldots F_{\sigma_0} \). By Corollary 1, \( F_{\sigma} \) has the property that \( F_{\sigma} \neq 0 \) if and only if \( \sigma \in L(\mathcal{M}) \); furthermore, there is at most one entry “1” in each column of \( F_{\sigma} \) with other entries being “0”. Then, by the definition of the norm \( \|\cdot\| \) in (30), it is easy to see that \( \|\Phi_{\sigma}\| = \|A_{\sigma}\| \) \( \) when \( \sigma \in L(\mathcal{M}) \), and \( \|\Phi_{\sigma}\| = 0 \) when \( \sigma \notin L(\mathcal{M}) \). Hence, it holds that \( \rho_k(\mathcal{A}_M) = \rho_k(\mathcal{A}, \mathcal{M}) \) for any \( k \in \mathbb{Z}_{\geq 0} \). Since \( \rho(\mathcal{A}_M) = \limsup_{k \to \infty} \rho_k(\mathcal{A}_M)^{1/k} \) and \( \rho(\mathcal{A}, \mathcal{M}) = \limsup_{k \to \infty} \rho_k(\mathcal{A}, \mathcal{M})^{1/k} \), it follows that

\[
\rho(\mathcal{A}_M) = \rho(\mathcal{A}, \mathcal{M}).
\]

(32)

Recalling (3) and (7), \( \rho_k(\mathcal{A}_M) \) is defined as \( \rho_k(\mathcal{A}_M) = \max_{\sigma \in [m]^k} \rho(\Phi_{\sigma}) \) and \( \bar{\rho}_k(\mathcal{A}, \mathcal{M}) \) is defined as \( \bar{\rho}_k(\mathcal{A}, \mathcal{M}) = \max_{\sigma \in [m]^k, \sigma \in L(\mathcal{M})} \rho(\mathcal{A}_\sigma) \). For any input sequence \( \sigma, \mathcal{F}_\sigma \) is a square matrix and has at most one entry “1” in each column with other entries being “0”. Then, it is easy to see that \( \mathcal{F}_\sigma \) is similar to a block diagonal matrix where the block matrix can be either a block zero matrix, or a strictly lower triangular matrix, or a permutation matrix. In other words, \( \mathcal{F}_\sigma \) is similar to a matrix \( \mathcal{K} := diag(K_1, K_2, K_3) \), denoted as \( \mathcal{F}_\sigma \sim_s \mathcal{K} \), where \( K_1 = \mathbf{0}, K_2 \) is a strictly lower triangular matrix, \( K_3 \) is a permutation matrix, but \( K_1, K_2, K_3 \) may not be present. Define \( L^{(per)}(\mathcal{M}) = \{ \sigma : \sigma \in L(\mathcal{M}), \mathcal{F}_\sigma \sim_s diag(K_1, K_2, K_3) \) where \( K_3 \) is present\}, where (per) stands for “permutation
matrix”. We claim that 1) the spectral radius of \( F \) is equal to 1 if and only if \( \sigma \in L^{(per)}(\mathcal{M}) \); 2) for any \( k_0 \in \mathbb{Z}_{>0} \), there exist some \( k_0' > k_0 \) and \( \sigma \in L^{(per)}(\mathcal{M}) \) such that \( |\sigma| = k_0' \). Since the eigenvalues of a permutation matrix lie on the unit circle, the absolute value of the eigenvalues of a permutation matrix are equal to 1. Hence, given \( F \) where \( \sigma \in L^{(per)}(\mathcal{M}) \), the absolute value of its eigenvalues are either equal to 1 or 0, where the latter case happens when \( F \sim_{\sigma} \text{diag}(K_1, K_2, K_3) \) with \( K_1 \) or \( K_2 \) present; given \( F \) where \( \sigma \in L(\mathcal{M}), \sigma \notin L^{(per)}(\mathcal{M}) \), all of its eigenvalues are equal to 0. The first claim is thus proved. Given any \( k_0 \in \mathbb{Z}_{>0} \), there always exists some \( k_0' > k_0 \) such that \( \mathcal{M}(V, E) \) has a loop of length \( k_0' \), because \( \mathcal{M} \) is assumed to be alive. Hence, there exists some state \( q \in Q \) and input sequence \( \sigma \in L(\mathcal{M}) \) such that \( |\sigma| = k_0' \) and \( f(q, \sigma) = q \). This means that there is at least a “1” in the diagonal of \( F \), which implies that \( \sigma \in L^{(per)}(\mathcal{M}) \) by definition. The second claim is thus proved.

Define \( \bar{\rho}^{(per)}(\mathcal{A}, \mathcal{M}) = \max_{\sigma \in [m]^k, \sigma \in L^{(per)}(\mathcal{M})} \rho(\mathcal{A}_\sigma) \), and \( \rho^{(per)}(\mathcal{A}, \mathcal{M}) = \limsup_{k \to \infty} \bar{\rho}^{(per)}(\mathcal{A}, \mathcal{M})^{1/k} \). Recalling the second property of the Kronecker product in Section 2, it is easy to see that \( \bar{\rho}_k(\mathcal{A}_\mathcal{M}) = \bar{\rho}^{(per)}(\mathcal{A}, \mathcal{M}) \). Then it follows that \( \rho(\mathcal{A}_\mathcal{M}) = \bar{\rho}^{(per)}(\mathcal{A}, \mathcal{M}) \). By the Berger-Wang Theorem, it holds that \( \rho(\mathcal{A}_\mathcal{M}) = \bar{\rho}(\mathcal{A}_\mathcal{M}) \).

Therefore, \( \bar{\rho}(\mathcal{A}, \mathcal{M}) = \rho(\mathcal{A}, \mathcal{M}) \). Since \( L^{(per)}(\mathcal{M}) \subseteq L(\mathcal{M}) \), it holds that \( \rho^{(per)}(\mathcal{A}, \mathcal{M}) \leq \rho(\mathcal{A}, \mathcal{M}) \). By definition, it holds that \( \bar{\rho}(\mathcal{A}, \mathcal{M}) \leq \rho(\mathcal{A}, \mathcal{M}) \). Therefore, we have \( \rho(\mathcal{A}, \mathcal{M}) = \rho^{(per)}(\mathcal{A}, \mathcal{M}) \).

Combing (32), (33) and (34) completes the proof.

The importance of Theorem 2 lies in that the problem of approximating the CJSR/CGSR of \( S(\mathcal{A}, \mathcal{M}) \) can be converted into the problem of approximating the JSR/GSR of its lifted system \( S(\mathcal{A}_\mathcal{M}) \), for which many off-the-shelf algorithms exist. A summary of different algorithms to compute the JSR/GSR of the arbitrary switching system can be found in [23] where a Matlab toolbox was also provided. In the following example, we will use the example given in [20] to show the effectiveness of the lifting method proposed above.

**Example 3.** Consider the constrained switching system \( S(\mathcal{A}, \mathcal{M}) \) where the set \( \mathcal{A} \) is given in Example 1 and the DFA \( \mathcal{M} \) is given in Example 2. For \( i = 1, \ldots, 4 \), calculate the matrices \( \Phi_i = F_i \otimes A_i \) where \( F_i \) and \( A_i \) are given in Example 1 and Example 2 respectively. Then define the set of lifted matrices \( \mathcal{A}_\mathcal{M} = \{ \Phi_1, \Phi_2, \Phi_3, \Phi_4 \} \). We use the JSR toolbox in [23] to approximate the value of \( \rho(\mathcal{A}_\mathcal{M}) \), or equivalently, \( \rho(\mathcal{A}, \mathcal{M}) \), by the conclusion of Theorem 2. In a computer with 3.5GHz CPU and 16GB memory, it took about 13.7 seconds for the jsr function to return

\[
\rho(\mathcal{A}, \mathcal{M}) \leq 0.974817295
\]

where the Gripenberg’s algorithm and the conitope algorithm are utilized [13, 10].

To compute the CJSR of the constrained switching system, [24] proposed the multinorm method where several algebraic lifting methods (e.g. the T-product
lift, the $M$-path-dependent lift that will be discussed in Section 5 were also combined to improve the estimation accuracy. Using the toolbox provided in [20] and fixing $T = 8$, it takes about 581 seconds (on the same computer as above) for the $T$-product lift to obtain the bounds

$$0.9289 \leq \rho(\mathcal{A}, \mathcal{M}) \leq 0.9761,$$

while by fixing $M = 7$, it takes about 1156 seconds for the $M$-path-dependent lift to obtain the bounds

$$0.9277 \leq \rho(\mathcal{A}, \mathcal{M}) \leq 0.9748.$$

It is obvious that our lifting method returns a more accurate approximation of $\rho(\mathcal{A}, \mathcal{M})$ in a much shorter time.

5 Discussion

The lifting method proposed in the preceding sections not only has the computational advantages in approximating the CJSR, but also has a clear interpretation as shown by the STP-based matrix expressions for the arbitrary/constrained switching systems. In this section, we will show further that the STP-based matrix approach can provide insight to some other lift-based or norm-based methods that are used for approximating CJSR.

5.1 The Kronecker Lift

Two lifting methods that are both based on the Kronecker product were proposed for the constrained switching system in [14] and [24]. Given a finite set of matrices $\mathcal{A} = \{A_1, \ldots, A_m\}$ where $A_i \in \mathbb{R}^{n \times n}$, $i \in [m]$, a DFA $\mathcal{M}(V, E)$ where $V = \{v_1, \ldots, v_T\}$, the lifting of $\mathcal{A}$ and $\mathcal{M}$ proposed in [24] is defined as a finite set of matrices $\tilde{\mathcal{A}}_{\mathcal{M}} := \{A_{(v_i, v_j, s)} : (v_i, v_j, s) \in E\}$ where $A_{(v_i, v_j, s)} := (\delta_{ij}^s)(\delta_i^j)^T \odot A_{s}$, $\odot$ is the Kronecker product. It is not hard to see that the lifting matrix $\Phi_s$ defined in (27) has the same size as $A_{(v_i, v_j, s)}$, and it holds that $F_s = \sum_{(i,j, (v_i, v_j, s) \in E)} \delta_i^s(\delta_j)^T$ for any $s \in [m]$, where the matrix $F_s$ is defined in (22). Therefore, the matrix $\Phi_s$ satisfies $\Phi_s = \sum_{(i,j, (v_i, v_j, s) \in E)} A_{(v_i, v_j, s)}$ for any $s \in [m]$. It is also clear that the number of matrices in the set $\tilde{\mathcal{A}}_{\mathcal{M}}$, which is defined in (29), is no larger than the number of matrices in the set $\tilde{\mathcal{A}}_{\mathcal{M}}$: $|\tilde{\mathcal{A}}_{\mathcal{M}}| \leq |\tilde{\mathcal{A}}_{\mathcal{M}}|$.

Given a finite set of matrices $\mathcal{A} = \{A_1, \ldots, A_m\}$ where $A_i \in \mathbb{R}^{n \times n}$, $i \in [m]$ and a matrix $\Omega = (\omega_{ij}) \in \mathbb{R}^{m \times m}$ where $\omega_{ij} \in \{0, 1\}$, the matrix product $A_{\sigma_k} \ldots A_1$ for $k \geq 2$ is called Markovian if each pair of indices $\{\sigma_i, \sigma_{i+1}\}$ is $\Omega$-admissible, i.e., $\omega_{\sigma_{i+1}\sigma_i} = 1$ for all $i \in [k-1]$. Then the Markovian joint spectral radius of $\mathcal{A}$ and $\Omega$ is defined as $\rho(\mathcal{A}, \Omega) := \limsup_{k \to \infty} \rho_k(\mathcal{A}, \Omega)^{1/k}$ where $\rho_k(\mathcal{A}, \Omega) = \max\{\|A_{\sigma_k} \ldots A_1\| : \omega_{\sigma_{i+1}\sigma_i} = 1, \forall i \in [k-1]\}$, and the Markovian generalized spectral radius of $\mathcal{A}$ and $\Omega$ is defined as $\overline{\rho}(\mathcal{A}, \Omega) := \limsup_{k \to \infty} \rho_k(\mathcal{A}, \Omega)^{1/k}$ where $\rho_k(\mathcal{A}, \Omega) = \max\{\rho(A_{\sigma_k} \ldots A_1) : \omega_{\sigma_{i+1}\sigma_i} = 1, \forall i \in [k-1]\}$.

The $\Omega$-lift of $\mathcal{A}$ defined in [14] is a finite set of matrices $\tilde{\mathcal{A}}_{\Omega} := \{\Omega \otimes A_i : i \in [m]\}$ where $\Omega$ is the $i$-th column of $\Omega$. By using the lifted system $\tilde{\mathcal{A}}_{\Omega}$, the Markovian analog of the Berger-Wang formula was proved in [14]: the Markovian joint spectral radius of $\mathcal{A}$ and $\Omega$ is equivalent to the Markovian generalized spectral radius of $\mathcal{A}$ and $\Omega$, which is also equivalent to the joint/generalized
spectral radius of $A_0$. We point out that the switching sequences for $A_M$ and $A_0$ are constrained by a DFA $M$ and a matrix $\Omega$, respectively. The matrices in $A_M$ are defined in a similar manner as in $A_0$; however, the STP-based matrix formulation provides a clear interpretation for $A_M$, as $\Phi$ is the transition structure matrix of the constrained switching system $S(A, M)$.

5.2 The Multinorm Method

A multinorm-based method was proposed to approximate $\rho(A, M)$ in [20]. Given a finite set of matrices $A = \{A_1, \ldots, A_m\}$ where $A_i \in \mathbb{R}_n^{n \times n}$, $i \in [m]$ and a DFA $M(V, E)$ where $|V| = \ell$, the multinorm of $S(A, M)$ associates a different norm for each node of the DFA $M$. Specifically, a multinorm $H$ for the system $S(A, M)$ is a set of $\ell$ norms $H = \{||v||, v \in V\}$. The value $\gamma^*(H)$ of a multinorm is defined as $\gamma^*(H) = \min_{x} \{\|Ax\|_v, \forall x \in \mathbb{R}^n, (v, w, \sigma) \in E\}$.

The multinorm of $S(A, M)$ can be used to bound $\rho(A, M)$ as it was shown that $\rho(A, M)$ is the infimum of $\gamma^*(H)$ for all the possible multinorm of $S(A, M)$, i.e., $\rho(A, M) = \inf_{H} (\gamma^*(H) : H \text{ is a multinorm for } S(A, M))$. One can use John’s Ellipsoid Theorem to compute lower and upper bounds of $\rho(A, M)$ by considering the quadratic multinorm and solving a quasi-convex optimization. Specifically, given a system $S(A, M)$, the value $\gamma_*(S)$ such that

$$\gamma_*(S) = \inf_{\gamma \in R, Q \in R^{n \times n}, v \in V} \gamma$$

s.t. $\gamma^2Q_v - A_v^\top Q_v A_\sigma \succeq 0, \forall (v, w, \sigma) \in E,$

$$Q_v \succeq 0, \forall v \in V,$$

satisfies $\rho(A, M) \leq \gamma_*(S) \leq \sqrt{n} \rho(A, M)$.

The STP formulation of $S(A, M)$ provides a clear interpretation to [35]. Recall the matrix expression for $S(A, M)$ in [20] and the special block structure of $\xi$. The quadratic multinorm of $S(A, M)$ can be defined as $\xi^\top Q\xi$ where the matrix $Q$ satisfies $Q := \text{diag}(Q_1, Q_2, \ldots, Q_{\ell}) \succ 0$ for some $Q_1, \ldots, Q_{\ell} \in \mathbb{R}_n^{n \times n}$. Then it is not hard to see that the optimization [35] is equivalent to

$$\gamma_*(S) = \inf_{\gamma \in R, Q \in R^{n \times n}, i \in [m]} \gamma$$

s.t. $\gamma^2Q - T(\Phi_i^\top Q \Phi_i) \succeq 0, i \in [m],$

$$Q \succeq 0,$$

where $\Phi_i$ is given in [27], $T(M) = M \odot (I_\ell \otimes 1_n^\top 1_n)$, $\odot$ is the Hadamard product.

5.3 The T-Product Lift

The so-called T-product lifting was proposed to increase the accuracy of the estimation of $\rho(A, M)$ [20]. Specifically, given a finite set of matrices $A = \{A_1, \ldots, A_m\}$ where $A_i \in \mathbb{R}_n^{n \times n}$, $i \in [m]$, a DFA $M(V, E)$ where $|V| = \ell$, and a positive integer $T$, the T-product of $S(TA, M)$, denoted $ST(A, M)$, is a constrained switching system on a DFA $M'(V', E')$ and a finite set of matrices $A'$ defined as follows: 1) $M'$ has the same set of nodes as $M$ (i.e., $V' = V$). For each path $p$ in $M$ that has length $T$ and connects $v$ and $w$, it is associated with an edge $e = (v, w, \sigma_1 \ldots \sigma_T)$ in $E'$. The label on this edge is a concatenation of those across the path $p$. 2) The set of matrices $A'$ is the set of all products
of size $T$ of matrices in $\mathcal{A}$ that are accepted by $\mathcal{M}$. For a label $\sigma_1 \ldots \sigma_T$ of the lifted system, $A_{\sigma_1} \ldots A_{\sigma_T} = A_{\sigma_T} \ldots A_{\sigma_1} \in \mathcal{A}'$.

An arbitrarily accurate estimate of $\rho(\mathcal{A}, \mathcal{M})$ can be calculated by the inequality $\gamma^*(\mathcal{S}^T) \leq n^2 \rho(\mathcal{A}, \mathcal{M})$, where $\gamma^*(\mathcal{S}^T)$ is the optimal value obtained by applying the optimization (35) to the lifted system $\mathcal{S}^T(\mathcal{A}, \mathcal{M})$ [20]. The DFA $\mathcal{M}'$ in $\mathcal{S}^T(\mathcal{A}, \mathcal{M})$ describes the evolution of $\mathcal{M}$ at times $iT$ for $i = 0, 1, 2, \ldots$, which can be easily constructed by using the STP formulation. Recalling the dynamics of $\mathcal{M}$ in (24), given $T \geq 2$, define $\tilde{q}(k) = q(Tk)$ and $\tilde{\sigma}(k) = \otimes_{j=1}^T \sigma(T(k+1)-j)$ for $k \in \mathbb{Z}_{\geq 0}$. Then it follows that $\tilde{q}(k+1) = q(T(k+1)) = F^T \otimes_{i=1}^{T-1} (I_{m_i} \otimes F) \otimes_{j=1}^T \sigma(T(k+1)-j) \otimes q(kT) = F_T \otimes \tilde{\sigma}(k) \otimes \tilde{q}(k)$, where $F_T := F^T \otimes_{i=1}^{T-1} (I_{m_i} \otimes F)$ with $F$ the transition structure matrix of $\mathcal{M}$. The DFA $\mathcal{M}'$ in $\mathcal{S}^T(\mathcal{A}, \mathcal{M})$ can be thus obtained from its transition structure matrix $F_T$. Recalling the dynamics of $\mathcal{S}(\mathcal{A})$ in (14) where $H$ is defined in (13), the matrix $\tilde{H}_T$ can be calculated using equation (17). Partitioning the matrix $\tilde{H}_T$ into $m^T$ sub-matrices $\tilde{H}_{Ti} \in \mathbb{R}^{n \times n}$ as in (18), then $\mathcal{A}'$ is just the set consisting of the sub-matrices $\tilde{H}_{Ti}$.

6 Conclusion

In this paper, we proposed a uniform, matrix-based formulation for the arbitrary switching system and the constrained switching system using the semi-tensor product of matrices, where the matrix expression of a constrained switching system can be seen as the matrix expression of a lifted arbitrary switching system. We proved that the constrained joint/generalized spectral radius of a constrained switching system is equivalent to the joint/generalized spectral radius of the lifted arbitrary switching system. Therefore, many off-the-shelf algorithms for approximating the joint/generalized spectral radius can now be used to approximate the constrained joint/generalized spectral radius. In future work, we plan to develop more efficient algorithms for approximating CJSR by incorporating the proposed lifting method and other lifting methods in the literature.

References

[1] A. A. Ahmadi, R. M. Jungers, P. A. Parrilo, and M. Roozbehani. Joint spectral radius and path-complete graph Lyapunov functions. SIAM Journal on Control and Optimization, 52(1):687–717, 2014.

[2] M. A. Berger and Y. Wang. Bounded semigroups of matrices. Linear Algebra and its Applications, 166:21–27, 1992.

[3] V. D. Blondel and Y. Nesterov. Computationally efficient approximations of the joint spectral radius. SIAM Journal on Matrix Analysis and Applications, 27(1):256–272, 2005.

[4] V. D. Blondel and J. N. Tsitsiklis. The boundedness of all products of a pair of matrices is undecidable. Systems & Control Letters, 41(2):135–140, 2000.
[5] D. Cheng and H. Qi. Controllability and observability of boolean control networks. *Automatica*, 45(7):1659–1667, 2009.

[6] D. Cheng and H. Qi. A linear representation of dynamics of boolean networks. *IEEE Transactions on Automatic Control*, 55(10):2251–2258, 2010.

[7] D. Cheng, H. Qi, and Y. Zhao. *An introduction to semi-tensor product of matrices and its applications*. World Scientific, 2012.

[8] X. Dai. A gelfand-type spectral radius formula and stability of linear constrained switching systems. *Linear Algebra and its Applications*, 436(5):1099–1113, 2012.

[9] I. Daubechies and J. C. Lagarias. Sets of matrices all infinite products of which converge. *Linear algebra and its applications*, 161:227–263, 1992.

[10] G. Gripenberg. Computing the joint spectral radius. *Linear Algebra and its Applications*, 234:43–60, 1996.

[11] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, 1990.

[12] R. Jungers. *The joint spectral radius: theory and applications*, volume 385. Springer Science & Business Media, 2009.

[13] R. M. Jungers, A. Cicone, and N. Guglielmi. Lifted polytope methods for computing the joint spectral radius. *SIAM Journal on Matrix Analysis and Applications*, 35(2):391–410, 2014.

[14] V. Kozyakin. The Berger-Wang formula for the Markovian joint spectral radius. *Linear Algebra and its Applications*, 448:315–328, 2014.

[15] B. Legat, R. M. Jungers, and P. A. Parrilo. Generating unstable trajectories for switched systems via dual sum-of-squares techniques. In *Proceedings of the 19th International Conference on Hybrid Systems: Computation and Control*, pages 51–60, 2016.

[16] D. Liberzon. *Switching in systems and control*. Springer Science & Business Media, 2012.

[17] H. Lin and P. J. Antsaklis. Stability and stabilizability of switched linear systems: a survey of recent results. *IEEE Transactions on Automatic control*, 54(2):308–322, 2009.

[18] M. Lothaire. *Algebraic combinatorics on words*, volume 90. Cambridge University Press, 2002.

[19] P. A. Parrilo and A. Jadbabaie. Approximation of the joint spectral radius using sum of squares. *Linear Algebra and its Applications*, 428(10):2385–2402, 2008.

[20] M. Philippe, R. Essick, G. E. Dullerud, and R. M. Jungers. Stability of discrete-time switching systems with constrained switching sequences. *Automatica*, 72:242–250, 2016.
[21] G.-C. Rota and W. Strang. A note on the joint spectral radius. *Proceedings of the Netherlands Academy*, 22:379–381, 1960.

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

[23] G. Vankeerberghen, J. Hendrickx, and R. M. Jungers. JSR: A toolbox to compute the joint spectral radius. In *Proceedings of the 17th international conference on Hybrid systems: computation and control*, pages 151–156, 2014.

[24] Y. Wang, N. Roohi, G. E. Dullerud, and M. Viswanathan. Stability analysis of switched linear systems defined by regular languages. *IEEE Transactions on Automatic Control*, 62(5):2568–2575, 2017.

[25] X. Xu and Y. Hong. Matrix expression and reachability analysis of finite automata. *Journal of Control Theory and Applications*, 10(2):210–215, 2012.

[26] X. Xu and Y. Hong. Matrix approach to model matching of asynchronous sequential machines. *IEEE Transactions on Automatic Control*, 58(11):2974–2979, 2013.

[27] X. Xu and Y. Hong. Observability analysis and observer design for finite automata via matrix approach. *IET Control Theory & Applications*, 7(12):1609–1615, 2013.