

# Switched Boolean Network Identification under Multiple Samples

Chunfeng Jiang<sup>1</sup>, Carmen Del Vecchio<sup>1</sup>, and Biao Wang<sup>2</sup>

**Abstract**—This paper investigates the challenging task of identifying switched Boolean networks (SBNs) with unknown initial subnetworks, leveraging the semi-tensor product tool. The inherent complexity arising from the lack of knowledge regarding the initial subnetwork necessitates a rigorous approach to identification. Initially, the observability of SBNs is discussed using a matrix-based method, providing foundational insights for subsequent identification analysis. We find the accurate one-to-one correspondence between states and outputs by taking time into account. Building on this correspondence and the observability property, the identification of state and output evolution rules is systematically addressed, leading to the proposition of a necessary and sufficient condition for successful SBN identification. Subsequently, an efficient algorithm is developed to implement the proposed identification approach. Finally, the theoretical findings are validated through an illustrative example.

## I. INTRODUCTION

The identification of gene regulatory networks (GRNs) holds significant importance in biological and medical research [1], [2]. GRNs are complex and precise regulatory systems within organisms that control gene expression and activity, ultimately determining cellular functions and characteristics. The introduction of Boolean networks (BNs) by Kauffman as a simplified model for simulating GRNs [3] has spurred considerable interest in the identification of BNs [4]. Notably, [5] presented the first systematic algorithmic approach to identifying GRNs based on a BN model. This was followed by [6], who proposed a straightforward algorithm for BN identification using state transition pairs, accompanied by a mathematical analysis. A significant advancement was made by [7], who leveraged the semi-tensor product (STP) of matrices to transform BN identification problems into algebraic problems involving structure matrices. This approach was later extended to Boolean control networks (BCNs) [8]. In recent years, the STP tool has facilitated numerous advances in the inference of B(C)Ns [9], [10], [11], [12]. Additionally, research has expanded to include the identification of delayed B(C)Ns and singular BCNs [13], [14].

This paper focuses on switched Boolean networks (SBNs), which are relevant not only to GRNs but also to a wide range of applications, including modeling eukaryotic cell division

cycles, bacteriophage genetic switches, and regulatory mechanisms in biological systems [15], [16]. For example, the eukaryotic cell division cycle can be modeled as a series of distinct processes, each governed by a unique set of regulatory rules. Similarly, the genetic switch in bacteriophages, which toggles between lysis and lysogeny states, provides a compelling application for SBNs. Another example is the lac operon in *Escherichia coli*, where extracellular glucose and lactose act as switching signals, highlighting the versatility of SBNs in capturing complex regulatory interactions. Notably, a BCN can also be viewed as an SBN by encoding control inputs as switching signals [17]. While significant progress has been made in the study of SBNs [18], [19], [20], [21], the identification of SBNs remains an underexplored area.

The methodology presented in this paper builds on the STP tool, which has proven effective in studying BNs [22]. By converting the dynamic characteristics of B(C)Ns into matrix operations, the STP enables the transformation of nonlinear systems into linear systems with unique state and output matrices. This approach has facilitated the resolution of critical problems such as reconstructibility, observability, decoupling, and output tracking [17], [23], [24], [25], [26].

The identification problem addressed in this paper is inspired by the fourth identification case in [12], which considers scenarios where observed data may originate from different initial configurations. While [12] proposed conditions for the identifiability of BCNs, these conditions are not directly applicable to our problem. This is because the system under consideration (also studied in [19]) cannot be treated as a traditional BCN; each of its subnetworks may produce distinct outputs, and the initial subnetwork is unknown. These complexities make the identification problem more challenging. However, the conditions proposed in [12] provide valuable insights for addressing the problem from the perspective of observability.

Drawing on existing literature on observability [28], particularly the matrix-based method introduced by [24] to reduce timing complexity in the study of weak observability in BNs, this paper proposes a condition for determining weak observability in SBNs using an equivalent matrix-based approach. Building on this, we establish a necessary and sufficient condition for identifying SBNs. Our analysis reveals that an SBN is observable if and only if its state equations (i.e., its state structure matrix) can be identified. However, complete inference of the output structure matrix requires additional conditions. Furthermore, we develop an algorithm to facilitate the identification of the state structure matrix and partial identification of the output structure matrix.

The remainder of the paper is organized as follows:

\*This work is supported by the program PRIN of the Italian government CUP C53D23008320001. (Corresponding author: Carmen Del Vecchio)

<sup>1</sup>Chunfeng Jiang and Carmen Del Vecchio are with the Department of Engineering, University of Sannio, Benevento 82100, Italy. [chunf\\_jiang@163.com](mailto:chunf_jiang@163.com), [c.delvecchio@unisannio.it](mailto:c.delvecchio@unisannio.it)

<sup>2</sup>Biao Wang is with School of Management, Shandong University, Jinan 250100, P.R. China. [wangbiao@sdu.edu.cn](mailto:wangbiao@sdu.edu.cn)

Section 2 introduces the notations used in this work. Section 3 defines the identification problem and converts SBNs into equivalent algebraic forms. Section 4 presents the main results, including criteria for determining the observability of SBNs, conditions for inferring their state and output equations, and a methodology for achieving identification. Section 5 provides an illustrative example to demonstrate the results. Finally, Section 6 concludes the paper.

## II. PRELIMINARIES

Notations:

- $\mathbb{N}$ ,  $\mathbb{Z}_+$  and  $\mathbb{R}_{p \times q}$  represent the sets of all natural numbers, positive integers, and  $p \times q$  real matrices, respectively.
- $\mathcal{D} := \{1, 0\}$  and  $\mathcal{D}^n := \underbrace{\mathcal{D} \times \cdots \times \mathcal{D}}_n$ .
- $\text{Col}_i(A)$  is the  $i$ -th column of a  $p \times q$  matrix  $A$  and  $\text{Col}(A) := \{\text{Col}_i(A) \mid i = 1, 2, \dots, q\}$ .
- $\Delta_m := \{\delta_m^i \mid i = 1, 2, \dots, m\}$ , where  $\delta_m^i = \text{Col}_i(I_m)$ .
- $A \in \mathbb{R}_{p \times q}$  is called a logical matrix if  $\text{Col}(A) \subseteq \Delta_p$ .
- $\mathcal{L}_{p \times q}$  is the set of  $p \times q$  logical matrices. If  $A \in \mathcal{L}_{p \times q}$ , it can be expressed as  $A = [\delta_p^{i_1} \delta_p^{i_2} \cdots \delta_p^{i_q}]$  and denoted briefly by  $A = \delta_p[i_1 \ i_2 \ \cdots \ i_q]$ .
- $(A)_{i,j}$  is the  $(i, j)$ -th element of matrix  $A$ .
- $\otimes$  is the Kronecker product of matrices.
- $\text{lcm}(p, q)$  is the least common multiple of numbers  $p$  and  $q$ .
- $A^T$ : the transpose of matrix  $A$ .
- $\mathbf{1}_m := \underbrace{[1, 1, \dots, 1]^T}_m$ .

*Definition 1:* ([22]) The semi-tensor product (STP) of two matrices  $A \in \mathbb{R}_{p \times q}$  and  $B \in \mathbb{R}_{s \times t}$  is defined as  $A \ltimes B = (A \otimes I_{\frac{\alpha}{q}})(B \otimes I_{\frac{\alpha}{s}})$ , where  $\alpha = \text{lcm}(q, s)$ .

Note that STP serves as a generalization of the traditional matrix product, thus “ $\ltimes$ ” will be omitted.

*Lemma 1:* ([22]) Let  $f: \mathcal{D}^n \rightarrow \mathcal{D}$  be a Boolean function. There must exist a unique matrix  $M_f \in \mathcal{L}_{2 \times 2^n}$ , referred to as the structure matrix of  $f$ , such that in the vector form we have

$$f(x_1, x_2, \dots, x_n) = M_f \ltimes_{i=1}^n x_i,$$

where  $x_i \in \Delta_2, i \in \{1, 2, \dots, n\}$ .

## III. PROBLEM FORMULATION

The dynamics of a switched Boolean network (SBN) consisting of  $\mu$  subnetworks is described as:

$$\begin{cases} x_i(t+1) = f_i^{\sigma(t)}(x_1(t), x_2(t), \dots, x_n(t)), & i = 1, 2, \dots, n, \\ y_j(t) = h_j^{\sigma(t-1)}(x_1(t), x_2(t), \dots, x_n(t)), & j = 1, 2, \dots, p, \end{cases} \quad (1)$$

where  $\sigma(\cdot): \mathbb{N} \rightarrow \{1, 2, \dots, \mu\}$  is the switching signal,  $x_i \in \mathcal{D}$  for  $i = 1, 2, \dots, n$  represents the state node,  $y_j \in \mathcal{D}$  for  $j = 1, 2, \dots, p$  denotes the output node, and  $f_i^{\sigma(t)}: \mathcal{D}^n \rightarrow \mathcal{D}$  and  $h_j^{\sigma(t-1)}: \mathcal{D}^n \rightarrow \mathcal{D}$  are Boolean functions. Note that unlike traditional SBNs [17], [20], the initial activating subnetwork in system (1) is unknown, i.e., outputs at  $t = 0$  are undefined

and depend on the switching signal  $\sigma(t)$  at time  $t + 1$ . This version of SBN was proposed for the first time in [19].

Using Lemma 1, the system (1) can be equivalently expressed in algebraic form as:

$$\begin{cases} x(t+1) = L_{\sigma(t)}x(t), \\ y(t) = H_{\sigma(t-1)}x(t), \end{cases} \quad (2)$$

where  $x(t) = \ltimes_{i=1}^n x_i(t) \in \Delta_{2^n}$ ,  $y(t) = \ltimes_{j=1}^p y_j(t) \in \Delta_{2^p}$ , and  $L_k \in \mathcal{L}_{2^n \times 2^n}$  and  $H_k \in \mathcal{L}_{2^p \times 2^n}$  are the state and output structure matrices of the  $k$ -th subnetwork, respectively.

Define  $L := [L_1 \ L_2 \ \cdots \ L_\mu] \in \mathcal{L}_{2^n \times \mu 2^n}$  and  $H := [H_1 \ H_2 \ \cdots \ H_\mu] \in \mathcal{L}_{2^p \times \mu 2^n}$ . Then, the SBN (2) can be rewritten as:

$$\begin{cases} x(t+1) = L\sigma(t)x(t), \\ y(t) = H\sigma(t-1)x(t), \end{cases} \quad (3)$$

where  $\sigma(\cdot) \in \Delta_\mu$ . It is clear that  $L\sigma(t) = L_{\sigma(t)}$  and  $H\sigma(t-1) = H_{\sigma(t-1)}$ .

Following [12] we will make to following Assumption to ensure problem tractability.

*Assumption 1:* ([12]) The available switching signal data and observed data are sufficient to cover all possible switching-output pairs in the system.

Assumption 1 is reasonable in practical applications. For example, in disease studies, a large number of samples can be collected from multiple patients, and multiple samples are often obtained from a single patient to analyze the underlying disease evolutions. Based on this assumption, we define the identification problem as follows.

*Definition 2:* The SBN (3) is identifiable under multiple samples if the matrices  $L$  and  $H$  can be uniquely determined using  $m$  groups of switching-output data  $\{(\sigma^{j,i}(0), y^{j,i}(1)), (\sigma^{j,i}(1), y^{j,i}(2)), \dots, (\sigma^{j,i}(r-1), y^{j,i}(r)))\}$ , where  $i \in \{1, 2, \dots, m\}$ ,  $j \in \{1, 2, \dots, q\}$ , and  $r \in \mathbb{Z}_+$ . These data are generated from specific initial conditions.

This formulation implies that the data are collected from  $q$  initial states, with each state subjected to  $m$  distinct switching signal sequences. Note that  $x^{j,1}(0) = x^{j,2}(0) = \dots = x^{j,m}(0)$  for  $j \in \{1, 2, \dots, q\}$ . The length of the collected data may vary, leading to different values of  $r$  for each switching-output sequence.

*Remark 1:* If there exists a coordinate transformation between states  $w = Wx$ , where  $W \in \mathcal{L}_{2^n \times 2^n}$ , then the SBN (3) can be expressed as:

$$\begin{cases} w(t+1) = WL(I_\mu \otimes W^T)\sigma(t)w(t) := \hat{L}\sigma(t)w(t), \\ y(t) = H(I_\mu \otimes W^T)\sigma(t-1)w(t) := \hat{H}w(t). \end{cases}$$

The tuples  $(L, H)$  and  $(\hat{L}, \hat{H})$  belong to the same equivalence class, which cannot be distinguished by any switching-output data. Therefore, our goal is to identify an equivalence class of  $(L, H)$ .

## IV. MAIN RESULTS

In this section, we establish conditions for identifying SBNs and propose an algorithm for system identification. Building on the result from [12], where a BCN is shown

to be identifiable under multiple samples if and only if it is observable, we first characterize observability for SBNs.

### A. Observability of Switched Boolean Networks

While [27] introduced four definitions of observability for BCNs, we focus on weak observability for SBNs, hereafter termed simply "observability." Given the unknown initial subnetwork, we assume the output at  $t = 0$  is unavailable, leading to the following definition.

*Definition 3:* An SBN (3) is *observable* if for any two distinct initial states  $x_0$  and  $\bar{x}_0$ , there exists a switching signal sequence  $\Sigma(l) := \{\sigma(0), \sigma(1), \dots, \sigma(l-1)\}$  such that the output sequences satisfy

$$\{y(1), y(2), \dots, y(l)\} \neq \{\bar{y}(1), \bar{y}(2), \dots, \bar{y}(l)\}.$$

To analyze observability, we construct two matrices  $Q_l$  and  $G_l$  as:

$$Q_l = G_l := L_1^T H_1^T H_1 L_1 \circ L_2^T H_2^T H_2 L_2 \circ \dots \circ L_\mu^T H_\mu^T H_\mu L_\mu, \quad (4)$$

where  $\circ$  denotes the Hadamard product. For  $j > 1$ , recursively define:

$$Q_j = L_1^T Q_{j-1} L_1 \circ L_2^T Q_{j-1} L_2 \circ \dots \circ L_\mu^T Q_{j-1} L_\mu, \quad (5)$$

and

$$G_j = Q_j \circ Q_{j-1} \circ \dots \circ Q_2 \circ G_1 = Q_j \circ G_{j-1}. \quad (6)$$

*Lemma 2:* An SBN (3) is observable if and only if there exists an integer  $l^*$  such that

$$G_{l^*} = I_{2^n}.$$

*Proof:* (Necessity) Assume that the SBN (3) is observable. By Definition 3, for any two distinct initial states  $\delta_{2^n}^a \neq \delta_{2^n}^b$ , there exists a minimal integer  $d_{a,b} \geq 1$  and a switching sequence  $\Sigma(d_{a,b}) = \{\sigma(0), \sigma(1), \dots, \sigma(d_{a,b}-1)\}$  such that:

$$\begin{cases} H_{\sigma(k)} L_{\sigma(k)} \dots L_{\sigma(0)} \delta_{2^n}^a = H_{\sigma(k)} L_{\sigma(k)} \dots L_{\sigma(0)} \delta_{2^n}^b, \\ 0 \leq k < d_{a,b} - 1, \\ H_{\sigma(d_{a,b}-1)} L_{\sigma(d_{a,b}-1)} \dots L_{\sigma(0)} \delta_{2^n}^a \neq \\ H_{\sigma(d_{a,b}-1)} L_{\sigma(d_{a,b}-1)} \dots L_{\sigma(0)} \delta_{2^n}^b. \end{cases}$$

The recursive construction of the matrices  $G_j$  in (4)–(6) encodes the distinguishability of state pairs:

- $(G_j)_{a,b} = 1$  if and only if the states  $\delta_{2^n}^a$  and  $\delta_{2^n}^b$  produce identical output sequences for all switching sequences of length  $j$ .
- $(G_j)_{a,b} = 0$  if there exists at least one switching sequence of length  $j$  that distinguishes  $\delta_{2^n}^a$  and  $\delta_{2^n}^b$ .

From the minimality of  $d_{a,b}$ , it follows that:

$$(G_{d_{a,b}-1})_{a,b} = 1 \quad \text{and} \quad (G_{d_{a,b}})_{a,b} = 0.$$

Let  $l^* = \max_{a \neq b} d_{a,b}$ . By the following construction:

- For all  $a \neq b$ ,  $(G_{l^*})_{a,b} = 0$ ;
- the identity  $H_{\sigma(k)} L_{\sigma(k)} \dots L_{\sigma(0)} \delta_{2^n}^a = H_{\sigma(k)} L_{\sigma(k)} \dots L_{\sigma(0)} \delta_{2^n}^a$  holds trivially for all  $k \in \mathbb{N}$ , thus  $(G_j)_{a,a} = 1$  for all  $j \in \mathbb{Z}_+$ ,

we conclude that  $G_{l^*} = I_{2^n}$ .

(Sufficiency) Assume  $G_{l^*} = I_{2^n}$ . We show that the SBN is observable by constructing a distinguishing switching sequence for any pair of distinct initial states.

For any  $a \neq b$ ,  $(G_{l^*})_{a,b} = 0$  implies that there exists a switching sequence  $\Sigma(l^*) = \{\sigma(0), \sigma(1), \dots, \sigma(l^*-1)\}$  such that:

$$H_{\sigma(l^*-1)} L_{\sigma(l^*-1)} \dots L_{\sigma(0)} \delta_{2^n}^a \neq H_{\sigma(l^*-1)} L_{\sigma(l^*-1)} \dots L_{\sigma(0)} \delta_{2^n}^b.$$

This ensures that the output sequences starting from  $\delta_{2^n}^a$  and  $\delta_{2^n}^b$  are distinct under the switching sequence  $\Sigma(l^*)$ . Therefore, the SBN is observable.  $\blacksquare$

### B. Identification of Switched Boolean Networks

Building on the observability property, we now address the identification problem for the SBN (3). While switching signals may be treated analogously to control inputs in BCNs, the identification methodology proposed in [12] is insufficient for SBNs due to two critical distinctions:

- The SBN comprises  $\mu$  distinct Boolean subnetworks with an *unknown initial activating subnetwork*.
- The output of a state may vary in distinct subnetworks significantly.

These characteristics introduce challenges beyond those encountered in traditional BCN frameworks, necessitating a specialized approach for identifying the structure matrices  $L$  and  $H$ .

In SBN (3), there are  $2^n(2^n - 1)$  distinct initial state pairs  $(\delta_{2^n}^a, \delta_{2^n}^b)$  where  $\delta_{2^n}^a \neq \delta_{2^n}^b$ . To eliminate redundancy –i.e., to eliminate the switching sequence for the pairs  $(\delta_{2^n}^a, \delta_{2^n}^b)$  and  $(\delta_{2^n}^b, \delta_{2^n}^a)$  that would lead to the same output sequence– we restrict attention to ordered pairs satisfying  $1 \leq a < b \leq 2^n$ . The number of required switching sequences reduces to  $\omega := 2^{n-1}(2^n - 1)$ . Let  $\Sigma_{(a,b)}(l_{a,b})$  denote a switching sequence of length  $l_{a,b}$  that distinguishes states  $\delta_{2^n}^a$  and  $\delta_{2^n}^b$ . For analytical tractability, we standardize all sequences to a uniform length:

$$l^* = \max_{a \neq b} \{l_{a,b}\},$$

extending shorter sequences as needed. These standardized sequences are then ordered via the bijective mapping:

$$e(a,b) = \begin{cases} b-a, & a=1, \\ (2^n - a/2)(a-1) + b-a, & \text{otherwise.} \end{cases} \quad (7)$$

Clearly,  $e(a,b) \in \{1, 2, \dots, \omega\}$ . The resulting the sequences

$$\Sigma^{e(a,b)}(l^*) = \{\sigma^{e(a,b)}(0), \dots, \sigma^{e(a,b)}(l^*-1)\} = \Sigma_{(a,b)}(l^*), \quad (8)$$

which called *O-switching sequences* here, guarantee observability of the SBN (3).

Let  $Y^{j,e} = \{y^{j,e}(1), y^{j,e}(2), \dots, y^{j,e}(l^*)\}$  denote the output sequence generated from initial state  $\delta_{2^n}^j$  under the  $e$ -th *O-switching sequence*  $\Sigma^e(l^*)$ . We define the *output sequence family* for each state as:

$$M^j := \{Y^{j,1}, Y^{j,2}, \dots, Y^{j,\omega}\}, \quad j \in \{1, 2, \dots, 2^n\}. \quad (9)$$

*Proposition 1:* For any distinct states  $\delta_{2^n}^a \neq \delta_{2^n}^b$ :

- 1)  $Y^{a,e(a,b)} \neq Y^{b,e(a,b)}$ ,

2)  $M^a \neq M^b$ .

This one-to-one correspondence between states and output families is a direct consequence of observability.

Indeed, we can collect abundant data samples, which most likely span the entire state space. This is to ensure that all possible switching-output data with respect to  $2^n$  distinct initial states can be gleaned. Combining this with the observability property, we establish:

*Theorem 1:* The state structure matrix  $L$  of SBN (3) is identifiable if and only if the system is observable.

*Proof:* (Sufficiency): Assume the SBN is observable. By Definition 3, there exist O-switching sequences that distinguish all pairs of distinct initial states. Collect switching-output data from  $2^n$  initial states  $\delta_{2^n}^j$ ,  $j \in \{1, 2, \dots, 2^n\}$ , under  $(\mu + 1)\omega$  switching sequences. For each initial state  $\delta_{2^n}^j$ , construct a series of switching signal sequences:

$$\begin{aligned} \Sigma^{j,e}(l^*) &= \{\sigma^{j,e}(0), \dots, \sigma^{j,e}(l^* - 1)\} = \Sigma^e(l^*), \\ \Sigma^{j,\omega+e}(l^* + 1) &= \{\delta_{\mu}^1, \Sigma^e(l^*)\}, \\ &\vdots \\ \Sigma^{j,\mu\omega+e}(l^* + 1) &= \{\delta_{\mu}^{\mu}, \Sigma^e(l^*)\}, e \in \{1, 2, \dots, \omega\}. \end{aligned} \quad (10)$$

The corresponding output sequences are:

$$\begin{aligned} Y^{j,e}(l^*) &= \{y^{j,e}(1), \dots, y^{j,e}(l^*)\}, \\ Y^{j,\omega+e}(l^* + 1) &= \{y^{j,\omega+e}(1), \dots, y^{j,\omega+e}(l^* + 1)\}, \\ &\vdots \\ Y^{j,\mu\omega+e}(l^* + 1) &= \{y^{j,\mu\omega+e}(1), \dots, y^{j,\mu\omega+e}(l^* + 1)\}. \end{aligned} \quad (11)$$

Truncate each sequence to its last  $l^*$  outputs, denoted  $\bar{Y}^{j,i}$ , and group them into families:

$$M_k^j = \{\bar{Y}^{j,(k-1)\omega+1}, \bar{Y}^{j,(k-1)\omega+2}, \dots, \bar{Y}^{j,k\omega}\}, \quad (12)$$

for  $k \in \{1, 2, \dots, \mu + 1\}$ .

Observability ensures that the collection of  $M_k^j$  contains exactly  $2^n$  distinct output families  $M^1, M^2, \dots, M^{\frac{k}{2^n}}$ . Define:

$$M^{j'} = \{Y^{j',1}, Y^{j',2}, \dots, Y^{j',\omega}\},$$

where  $M^{j'}$  corresponds to the output family generated from  $\delta_{2^n}^{j'}$ . Match observed data to these families via:

$$\begin{cases} \tilde{x}^{j,1}(0) = \delta_{2^n}^{j'}, & \text{if } M_1^j = M^{j'}, \\ \tilde{x}^{j,i\omega}(1) = \delta_{2^n}^{j'}, & \text{if } M_i^j = M^{j'}, \\ \tilde{y}^{j,i\omega}(1) = y^{j,i\omega}(1), \end{cases} \quad (13)$$

for  $j \in \{1, 2, \dots, 2^n\}$  and  $i \in \{2, 3, \dots, \mu + 1\}$ .

From the switching signals  $\sigma^{j,i\omega}(0)$ , we obtain

$$\begin{cases} \tilde{x}^{j,i\omega}(1) = L\sigma^{j,i\omega}(0)\tilde{x}^{j,1}(0), \\ \tilde{y}^{j,i\omega}(1) = H\sigma^{j,i\omega}(0)\tilde{x}^{j,i\omega}(1), \end{cases} \quad (14)$$

which uniquely identifies the structure matrices  $L$ . The bijection between states and output families ensures all columns of  $L$  are identifiable.

(Necessity) We prove necessity by contradiction. If the system is unobservable, two distinct states yield identical

outputs under all switching sequences, preventing unique identification of  $L$ . ■

Theorem 1 shows that  $L$  is fully identifiable under observability; in contrast the output matrix  $H$  requires additional conditions. This is due to fact that the second equation in (14) indicates that the identification of  $H$  relies on knowledge of state  $x(1)$  rather than  $x(0)$ . Therefore, identification of both  $L$  and  $H$  requires not only the observability of system (3) but also the additional condition, which is

$$L_i \mathbf{1}_{2^n} = \mathbf{1}_{2^n}, \quad \forall i \in \{1, \dots, \mu\}. \quad (15)$$

This ensures every state is reachable at  $t = 1$ . Let  $\mathcal{M}_k := \{M_k^j \mid j = 1, 2, \dots, 2^n\}$  with  $k \in \{2, 3, \dots, \mu + 1\}$  and  $\mathcal{M} := \{M^{j'} \mid j' = 1, 2, \dots, 2^n\}$ . From (13), we know that for the  $i$ -th subnetwork, if  $\mathcal{M}_{i+1} = \mathcal{M}$ , then  $x(1)$  can cover the whole state space, further, all columns of  $H_i$  can be determined according to (14).

*Theorem 2:* An SBN (3) is identifiable if and only if it is observable and  $\mathcal{M}_{i+1} = \mathcal{M}$  for all  $i \in \{1, \dots, \mu\}$ .

*Remark 2:* Existing methods for BCNs [12] are insufficient for SBNs due to asynchronous state-output evolution and varying subnetwork outputs. This work extends the results of [19] to address the identifiability. Furthermore, while [19] addresses reconstruction problems for SBN (3), it does not cover identification issues. This paper expands the scope of SBNs research by addressing these gaps.

Algorithm 1 identifies  $L$  and partial columns of  $H$ . The key steps include data collection, output family matching, and matrix determination.

Figure 1 visually illustrates the identification process described above: yellow circles denote states, while green circles represent outputs. There are  $2\omega$  groups of data, stemming from the same initial states  $x^{1,1}(0) = x^{1,2}(0) = x^{1,2\omega}(0) = \delta_{2^n}^1$ . And  $\sigma^{1,\omega+e}(0) = \delta_{\mu}^1$ ,  $e \in \{1, 2, \dots, \mu\}$ . Then, the outputs in red rectangle are the same. According to  $M_1^1 = M^1$ , we can identify  $\tilde{x}^{1,1}(0) = \delta_{2^n}^1$ . If  $M_2^1 \neq M_1^1$ , then  $M_2^1 = M^2$ . Thus,  $\tilde{x}^{1,2\omega}(1) = \delta_{2^n}^2$ . Eventually, based on (14),  $\text{Col}_1(L_1) = 2$  and  $\text{Col}_2(H_1) = y^{1,2\omega}(1)$  are identified.

## V. EXAMPLE

Consider an observable SBN with 2 subnetworks, 3 state nodes, 1 output node, and

$$\begin{cases} \hat{L} = \delta_8[3 \ 2 \ 4 \ 2 \ 8 \ 3 \ 6 \ 4 \ 1 \ 1 \ 7 \ 5 \ 5 \ 6 \ 7 \ 8], \\ \hat{H} = \delta_2[1 \ 2 \ 1 \ 1 \ 1 \ 2 \ 2 \ 1 \ 1 \ 2 \ 2 \ 1 \ 2 \ 2 \ 1 \ 2]. \end{cases} \quad (16)$$

We assume O-switching sequences of the system are

$$\Sigma^e = \begin{cases} \delta_2\{1, 2\}, e \in \{2, 3\} := E_1, \\ \delta_2\{2, 1\}, e \in \{4, 5, 9, 18\} := E_2, \\ \delta_2\{1, 1\}, e \in \{1, 2, \dots, 28\} \setminus (E_1 \cup E_2) := E_3. \end{cases} \quad (17)$$

---

**Algorithm 1** Identify  $L$  and some columns of  $H$  of SBN (3)
 

---

**Input:**  $s+2$  groups of switching signal data (10)

**Output:**  $L$  and  $H$

- 1: generate the  $(\mu+1)\omega$  groups of output data (11) and denote  $\bar{Y}^{j,i}$  as the output sequence consisted of the last  $l^*$  elements of sequence  $Y^{j,i}(l)$ , where  $i \in \{1, 2, \dots, (\mu+1)\omega\}$ . Group into families  $M_k^j$ .
  - 2: **Initialization** :  $M = \emptyset$ ,  $j' = 1$ ,
  - 3: **for** each  $j \in \{1, \dots, 2^n\}$  **do**
  - 4:   **for** each  $k \in \{1, \dots, \mu+1\}$  **do**
  - 5:     **if**  $M_k^j \notin M$  **then**
  - 6:        $M_k^j = M_k^j$ ,  $M \leftarrow M \cup \{M_k^j\}$ ,  $j' \leftarrow j' + 1$
  - 7:     **end if**
  - 8:   **end for**
  - 9: **end for**
  - 10: construct  $L$  and some columns of  $H$  using (13) and (14).
- 

Then, we construct the switching sequences according (17) and switching signal  $\delta_\mu^i$ :

$$\Sigma^{j,e}(2) = \begin{cases} \delta_2\{1, 2\}, & e \in E_1, \\ \delta_2\{2, 1\}, & e \in E_2, \\ \delta_2\{1, 1\}, & e \in E_3, \end{cases}$$

$$\Sigma^{j,28i+e}(3) = \begin{cases} \delta_2\{i, 1, 2\}, & e \in E_{3i+1}, \\ \delta_2\{i, 2, 1\}, & e \in E_{3i+2}, \\ \delta_2\{i, 1, 1\}, & e \in E_{3i+3}, \end{cases}$$

where  $i \in \{1, 2\}$ ,  $j \in \{1, 2, \dots, 8\}$ ,  $E_{3i+1} := \{28i + 23\}$ ,  $E_{3i+2} := \{28i + 4, 28i + 5, 28i + 9, 28i + 18\}$  and  $E_{3i+3} := \{28i + 1, 28i + 2, \dots, 28i + 28\} \setminus (E_{3i+1} \cup E_{3i+2})$ . Let the initial state  $x^{j,1}(0) = x^{j,2}(0) = \dots = x^{j,84}(0) = \delta_8^j$  and consider state  $\delta_8^1$  firstly. Then the output sequences with respect to state  $\delta_8^1$  are obtained:

$$Y^{1,e}(2) = \begin{cases} \delta_2\{1, 1\}, & e \in E_1, \\ \delta_2\{1, 1\}, & e \in E_2, \\ \delta_2\{1, 1\}, & e \in E_3, \end{cases}$$

$$Y^{1,28+e}(3) = \begin{cases} \delta_2\{1, 1, 2\}, & e \in E_4, \\ \delta_2\{1, 1, 2\}, & e \in E_5, \\ \delta_2\{1, 1, 2\}, & e \in E_6, \end{cases}$$

$$Y^{1,56+e}(3) = \begin{cases} \delta_2\{1, 1, 1\}, & e \in E_7, \\ \delta_2\{1, 1, 1\}, & e \in E_8, \\ \delta_2\{1, 1, 1\}, & e \in E_9. \end{cases}$$

Similarly, we can get the output sequences with respect to remaining states as

$$Y^{2,e}(2) = \begin{cases} \delta_2\{2, 1\}, & e \in E_1, \\ \delta_2\{1, 1\}, & e \in E_2, \\ \delta_2\{2, 2\}, & e \in E_3, \end{cases}$$

$$Y^{2,28+e}(3) = \begin{cases} \delta_2\{2, 2, 1\}, & e \in E_4, \\ \delta_2\{2, 1, 1\}, & e \in E_5, \\ \delta_2\{2, 2, 2\}, & e \in E_6, \end{cases}$$

$$Y^{2,56+e}(3) = \begin{cases} \delta_2\{1, 1, 1\}, & e \in E_7, \\ \delta_2\{1, 1, 1\}, & e \in E_8, \\ \delta_2\{1, 1, 1\}, & e \in E_9, \end{cases}$$

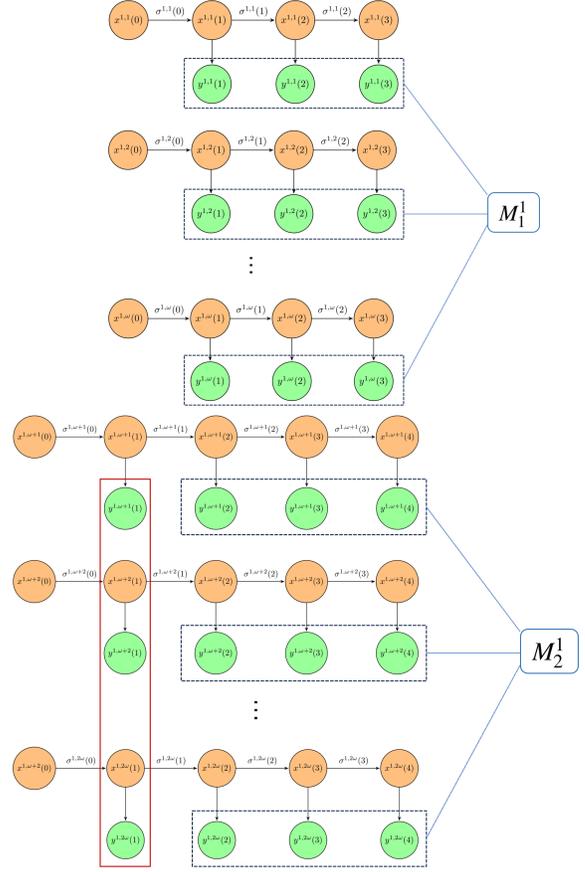


Fig. 1. Generate the first and second output sequence families  $M_1^1$  and  $M_2^1$

$$\begin{aligned} & \vdots \\ Y^{8,e}(2) &= \begin{cases} \delta_2\{1, 1, 1\}, & e \in E_1, \\ \delta_2\{2, 1, 1\}, & e \in E_2, \\ \delta_2\{1, 1, 2\}, & e \in E_3, \end{cases} \\ Y^{8,28+e}(3) &= \begin{cases} \delta_2\{1, 1, 1, 1\}, & e \in E_4, \\ \delta_2\{1, 1, 1, 1\}, & e \in E_5, \\ \delta_2\{1, 1, 2, 2\}, & e \in E_6, \end{cases} \\ Y^{8,56+e}(3) &= \begin{cases} \delta_2\{2, 1, 1, 1\}, & e \in E_7, \\ \delta_2\{2, 2, 1, 1\}, & e \in E_8, \\ \delta_2\{2, 1, 1, 2\}, & e \in E_9. \end{cases} \end{aligned}$$

From (12), we can get  $M_k^j$ . Then 8 distinct output sequence families are found as follows

$$\begin{aligned} M^1 &= M_1^1, M^2 = M_2^1, M^3 = M_1^2, M^4 = M_2^2, \\ M^5 &= M_3^3, M^6 = M_3^4, M^7 = M_2^5, M^8 = M_3^6. \end{aligned}$$

Denote that  $M^j$  is the output sequence groups of state  $\delta_{2^n}^j$  under switching sequences (17). Naturally, we can find the evolutionary trajectory of each state under switching signal  $\delta_2^1$  and  $\delta_2^2$ , respectively,

$$\begin{aligned} \delta_8^2 &= L\delta_2^1\delta_8^1, & \delta_8^1 &= L\delta_2^2\delta_8^1, & \delta_8^4 &= L\delta_2^1\delta_8^2, & \delta_8^5 &= L\delta_2^2\delta_8^2, \\ \delta_8^3 &= L\delta_2^1\delta_8^3, & \delta_8^1 &= L\delta_2^2\delta_8^3, & \delta_8^3 &= L\delta_2^1\delta_8^4, & \delta_8^6 &= L\delta_2^2\delta_8^4, \\ \delta_8^8 &= L\delta_2^1\delta_8^5, & \delta_8^5 &= L\delta_2^2\delta_8^5, & \delta_8^7 &= L\delta_2^1\delta_8^6, & \delta_8^6 &= L\delta_2^2\delta_8^6, \\ \delta_8^4 &= L\delta_2^1\delta_8^7, & \delta_8^7 &= L\delta_2^2\delta_8^7, & \delta_2^2 &= H\delta_2^1\delta_8^8, & \delta_2^8 &= H\delta_2^2\delta_8^8, \end{aligned}$$

At the same time, we know that

$$\begin{aligned}\delta_2^1 &= H\delta_2^2\delta_8^1, & \delta_2^1 &= H\delta_2^1\delta_8^2, \\ \delta_2^2 &= H\delta_2^1\delta_8^3, & \delta_2^1 &= H\delta_2^1\delta_8^4, \\ \delta_2^1 &= H\delta_2^2\delta_8^5, & \delta_2^2 &= H\delta_2^2\delta_8^6, \\ \delta_2^1 &= H\delta_2^1\delta_8^7, & \delta_2^2 &= H\delta_2^2\delta_8^7, \\ \delta_2^2 &= H\delta_2^1\delta_8^8, & \delta_2^1 &= H\delta_2^2\delta_8^8,\end{aligned}$$

Thus, based on the final step of Algorithm 1, the matrix  $L$  and some columns of  $H$  is identified as

$$\begin{cases} L = \delta_8[2 & 4 & 3 & 3 & 8 & 7 & 4 & 2 & 1 & 5 & 1 & 6 & 5 & 6 & 7 & 8], \\ H = \delta_2[* & 1 & 2 & 1 & * & * & 1 & 2 & 1 & * & * & * & 1 & 2 & 2 & 1], \end{cases}$$

where  $*$  is unknown number.

In fact, the equivalent class is identified by our method, and the coordinate transformation is  $x(t) = \delta_8[1 \ 3 \ 2 \ 4 \ 6 \ 8 \ 5 \ 7]\hat{x}(t)$ , where  $\hat{x}(t)$  is the state of the original system (16).

## VI. CONCLUSIONS

This paper addressed the identification of SBNs with unknown initial subnetworks using the STP tool. We established a matrix-based criterion for SBN observability and proved that the state structure matrix  $L$  is identifiable if and only if the system is observable. While  $L$  can be fully identified, complete inference of the output structure matrix  $H$  requires additional conditions. An efficient algorithm was proposed to identify  $L$  and partially identify  $H$ , validated through an illustrative example. Note that the computational complexity increase exponentially with the increase of nodes, thus the method is suitable to identify small scale SBNs.

Future research could focus on improving the scalability of large-scale networks and developing new approaches to identify such complex networks. Additionally, observational data are frequently affected by noise in practical applications due to constraints in sensing or communication. Thus, extending the current framework to handle observation noise and missing data is crucial for enhancing its robustness. Incorporating uncertainty and stochastic behavior, integrating machine learning, and applying the framework to probabilistic BN are also promising directions. Experimental validation with real biological data would further assess the practical utility of the proposed methods. These advancements could enhance the understanding and control of complex biological systems.

## REFERENCES

- [1] Y. Yuan, and Z. Bar-Joseph, Deep learning for inferring gene relationships from single-cell expression data, *Proceedings of the National Academy of Sciences*, 116(52), pp. 27151-27158, 2019.
- [2] P. Badia-i-Mompel, L. Wessels, S. Müller-Dott, R. Trimbour, R. Ramirez Flores, R. Argelaguet, and J. Saez-Rodriguez, Gene regulatory network inference in the era of single-cell multi-omics, *Nature Reviews Genetics*, 24(11), pp. 739-754, 2023.
- [3] S. Kauffman, Metabolic stability and epigenesis in randomly constructed genetic nets, *Journal of Theoretical Biology*, 22(3), pp. 437-467, 1969.
- [4] M. Ostrowski, L. Paulevé, T. Schaub, A. Siegel, and C. Guziolowski, Boolean network identification from perturbation time series data combining dynamics abstraction and logic programming, *Biosystems*, 149, pp. 139-153, 2016.

- [5] T. Akutsu, S. Kuhara, O. Maruyama, and S. Miyano, Identification of gene regulatory networks by strategic gene disruptions and gene overexpressions, *SODA*, 98, pp. 695-702, 1998.
- [6] T. Akutsu, S. Miyano, and S. Kuhara, Identification of genetic networks from a small number of gene expression patterns under the Boolean network model, *Proceedings of the Pacific Symposium on Biocomputing*, Mauna Lani, 4, pp. 17-28, 1999.
- [7] D. Cheng, H. Qi, Z. Li, Model construction of Boolean network via observed data, *IEEE Transactions on Neural Networks*, 22(4), pp. 525-536, 2011.
- [8] D. Cheng, Y. Zhao, Identification of Boolean control networks, *Automatica*, 47(4), pp. 702-710, 2011.
- [9] E. Fornasini, and M. Valcher, Identification problems for Boolean networks and Boolean control networks, *IFAC Proceedings Volumes*, 47(3), pp. 5399-5404, 2014.
- [10] X. Zhang, H. Han, and W. Zhang, Identification of Boolean networks using premined network topology information, *IEEE transactions on Neural Networks and Learning Systems*, 28(22), pp. 464-469, 2017.
- [11] T. Leifeld, Z. Zhang, and P. Zhang, Identification of Boolean network models from time series data incorporating prior knowledge, *Frontiers in Physiology*, 9, pp. 359680, 2018.
- [12] B. Wang, J. Feng, and D. Cheng, On identification of Boolean control networks, *SIAM Journal on Control and Optimization*, 60(3), pp. 1591-1612, 2022.
- [13] T. Mu, J. Feng, B. Wang, and Y. Jia, Identification of Boolean control networks with time delay, *ISA Transactions*, 144, pp. 113-123, 2024.
- [14] R. Zhao, J. Feng, Q. Meng, and B. Wang, Identification of a class of singular Boolean control networks, *Information Sciences*, 691, pp. 121627, 2025.
- [15] N. El-Farra, A. Gani and P. Christofides, Analysis of mode transitions in biological networks, *AIChE journal*, 51(8), pp. 2220-2234, 2005.
- [16] A. Veliz-Cuba, and B. Stigler, Boolean models can explain bistability in the lac operon, *Journal of Computational Biology*, 18(6), pp. 783-794, 2011.
- [17] E. Fornasini, and M. Valcher, Observability, reconstructibility and state observers of Boolean control networks, *IEEE Transactions on Automatic Control*, 58(6), pp. 1390-1401, 2013.
- [18] A. Yerudkar, C. Del Vecchio, L. Glielmo, Sampled-data set stabilization of switched Boolean control networks, *IFAC-PapersOnLine*, 53(2), pp.6139-6144, 2020.
- [19] Z. Gao, J. Feng, and B. Wang, On reconstructibility of switched Boolean control networks, *International Journal of Control*, 94(12), pp. 3339-3348, 2021.
- [20] L. Tong, J. Liang and H. Hu, Fault detectability of asynchronous switched Boolean networks: a set reachability approach, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 20(6), pp. 3876-3888, 2023.
- [21] C. Jiang, B. Wang, S. Fu, J. Zhao, and M. Sun, Observability of periodically switched Boolean control networks, *Journal of Systems Science and Complexity*, 36(3), pp. 985-1001, 2023.
- [22] D. Cheng, H. Qi, and Z. Li, *Analysis and Control of Boolean Networks: A semi-tensor Product Approach*, London: Springer, 2011.
- [23] Y. Guo, Observability of Boolean control networks using parallel extension and set reachability, *IEEE Transactions on Neural Networks and Learning Systems*, 29(12), pp. 6402-6408, 2018.
- [24] Y. Yu, M. Meng, J. Feng, and G. Chen, Observability criteria for Boolean networks, *IEEE Transactions on Automatic Control*, 67(11), pp. 6248-6254, 2022.
- [25] Y. Li, H. Lyu, J. Feng and A. Tayebi, New results on input-output decoupling of Boolean control networks, *Proceeding of 62nd IEEE Conference on Decision and Control*, Singapore, pp. 8334-8339, 2023.
- [26] A. Yerudkar, C. Del Vecchio, L. Glielmo, Output tracking control of probabilistic Boolean control networks, *2019 IEEE International Conference on Systems, Man and Cybernetics*, pp. 2109-2114, 2019.
- [27] K. Zhang and L. Zhang, Observability of Boolean control networks: a unified approach based on finite automata, *IEEE Transactions on Automatic Control*, 61(9), pp. 2733-2738, 2016.
- [28] K. Zhang, A survey on observability of Boolean control networks, *Control Theory and Technology*, 21(2), pp. 115-147, 2023.