

# Compositional and Abstraction-Based Approach for Synthesis of Edit Functions for Opacity Enforcement

Sahar Mohajerani, Yiding Ji and Stéphane Lafortune

**Abstract**—This paper develops a novel compositional and abstraction-based approach to synthesize edit functions for opacity enforcement in modular discrete event systems. Edit functions alter the output of the system by erasing or inserting events in order to obfuscate the outside intruder, whose goal is to infer the secrets of the system from its observation. We synthesize edit functions to solve the opacity enforcement problem in a modular setting, which significantly reduces the computational complexity compared with the monolithic approach. Two abstraction methods called opaque observation equivalence and opaque bisimulation are first employed to abstract the individual components of the modular system and their observers. Subsequently, we propose a method to transform the synthesis of edit functions to the calculation of modular supremal nonblocking supervisors. We show that the edit functions synthesized in this manner correctly solve the opacity enforcement problem.

## I. INTRODUCTION

Opacity characterizes whether the integrity of the secrets of a system can be preserved from the inference of an outside intruder, potentially with malicious purposes. The intruder is modeled as a passive observer with knowledge of the system's structure. A system is called opaque if the intruder is unable to infer the system's secrets from its observation.

Starting with [2], [3] in the computer science literature, opacity has been extensively studied, especially in the field of discrete event systems (DES), under multiple frameworks.

For finite state automaton models, various notions of opacity have been studied, e.g., language-based opacity [22], current-state opacity [34], initial-state opacity [36], K-step opacity [49] and infinite-step opacity [33]. Opacity has also been discussed in some other system models, like infinite state systems [6], modular systems [24] and Petri nets [42], [43]. Opacity under a special observer called Orwellian observer is discussed in [30] and opacity under powerful attackers is studied in [14]. A more recent work [51] investigates opacity for networked supervisory control systems. Furthermore, some works investigate opacity in stochastic settings, e.g., [1], [7], [21], [45]. Specifically, [52] presents a novel approach to tackle infinite-step and K-step opacity in stochastic DES. The survey paper [16] summarizes some recent results on opacity in DES.

When opacity does not hold, it is natural to study its enforcement [10]. One popular approach is supervisory control [8], [9], [35], [41], [48], where some behaviors of the system are disabled before they reveal the secrets. Another

widely-applied method is sensor activation [5], [50], [53], where the observability of events is dynamically changed.

Recently, a new enforcement method called insertion function was proposed in [46], which inserts fictitious events into the output of the system to obfuscate the intruder. The authors of [18] extended the method to study opacity enforcement under the assumption that the intruder may or may not know the implementation of insertion functions, while [19] discussed opacity enforcement by insertion functions under quantitative constraints. As a following work, [47] investigates a more general method called edit functions, which manipulate the output of the system by either inserting or erasing events. Then [17], [20] considers the case when the edit function's implementation is known to the intruder. As a summary and extension, [20] characterizes opacity enforcement by edit functions as a three-player game and proposes a novel information structure called three-player observer (TPO) to embed edit functions. A special TPO called the All Edit Structure (AES) is also introduced in [20] to characterize the edit constraints.

In this work, we elaborate the method in [20] to study opacity enforcement in a modular setting. Our motivation is as follows. To generate a three-player observer, the observer of the system needs to be calculated, which is potentially costly in computation. Furthermore, modern engineering systems usually contain multiple components that are synchronized and subject to malicious inference. In this sense, if we are to apply edit functions to enforce opacity, heavy computation is involved both from determining individual systems and synchronizing them, which may be potentially cumbersome.

To alleviate this issue, this paper applies a compositional and abstraction-based method to reduce the size of the modular system before calculating the All Edit Structure. *Bisimulation* and *observation equivalence* [25] are well-known methods to abstract the state space of an automaton, while they do not preserve opacity properties in general. As a variant, [54] proposes several innovative concepts termed opacity-preserving (bi)simulation relations to reduce the state space of the system in opacity verification. A compositional visible bisimulation equivalence method is discussed in [31] for abstraction-based opacity verification.

For abstraction, we introduce *opaque observation equivalence* and *opaque bisimulation*, which consider the secrecy status of states when merging them. In our framework, each individual system is abstracted using opaque observation equivalence. After that, the observer is calculated. Since abstraction reduces the size of the state space, the computational complexity of calculating the observer is lowered potentially. Next, opaque bisimulation is employed to the observer of each abstracted individual system, resulting in the smallest possible automaton for future discussion.

The work of the first author was supported by the Swedish Research Council. The work of the second and third authors was supported in part by US NSF grants CNS-1421122 and CNS-1738103.

Sahar Mohajerani, Yiding Ji and Stéphane Lafortune are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA. {saharm; jiyiding; stephane@umich.edu}

We further leverage some results from supervisor reduction and modular supervisory control theory to reduce the complexity of supervisor synthesis. There is a rich literature on both topics, see, e.g., [23], [28], [37], [39], [40]. The main idea is to convert the construction of the monolithic All Edit Structure to a modular supervisory control problem. Specifically, we first transfer each individual three-player observer (without considering edit constraints) to its automaton form and view the set of interacting automata as the “plant” to be controlled. Then we put the edit constraint as the specification, also in an automaton form. Afterwards, we perform modular supervisory control to synthesize a least restrictive and nonblocking modular supervisor. It is shown that all the traces accepted by the supervisor represent valid edit decisions contained in the monolithic AES. Compared with the conventional monolithic approach for supervisor synthesis [4], our compositional approach is more efficient in computation.

The presentation of this work is organized as follows. Section II gives a brief background introduction about the system model, supervisory control theory and edit functions. The general idea of the paper is presented in Sect. III as a flow chart. Section IV explains the abstraction methods and synchronization of three-player observers. Next, Section V transforms the calculation of the monolithic All Edit Structure to the calculation of a modular supervisor. Finally, some concluding remarks are given in Section VI.

A preliminary and partial version of this work appears in [26]. The current work improves [26] in the sense that [26] only considers abstraction methods to synthesize edit functions in a monolithic setting, while this work also takes synchronous composition into consideration and the edit functions are synthesized by a modular approach.

## II. MODELING FORMALISM AND BACKGROUND

### A. Events, Automata and their Composition

In this work, we consider discrete event systems modeled as deterministic or nondeterministic automata.

*Definition 1:* A (nondeterministic) finite-state automaton is a tuple  $G = \langle \Sigma, Q, \rightarrow, Q^0 \rangle$ , where  $\Sigma$  is a finite set of events,  $Q$  is a finite set of states,  $\rightarrow \subseteq Q \times \Sigma \times Q$  is the *state transition relation*, and  $Q^0 \subseteq Q$  is the set of *initial states*.  $G$  is *deterministic* if  $|Q^0| = 1$  and if  $x \xrightarrow{\sigma} y_1$  and  $x \xrightarrow{\sigma} y_2$  always implies that  $y_1 = y_2$ .

When state marking is considered, the above definition is extended to  $G = \langle \Sigma, Q, \rightarrow, Q^0, Q^m \rangle$ , where  $Q^m \subseteq Q$  is the set of *marked states*. In this paper, we identify marked states using gray shading in the figures.

We assume that the system is partially observed, thus the concepts of *observable* and *unobservable* events are introduced. Since the exact identity of unobservable events is irrelevant in our later discussion of opacity, they are uniformly represented by a special event  $\tau$ . The event  $\tau$  is never included in the alphabet  $\Sigma$ , unless explicitly mentioned. For this reason,  $\Sigma_\tau = \Sigma \cup \{\tau\}$  is used to represent the whole set of observable and unobservable events. Hereafter, nondeterministic automata may contain transitions labeled by  $\tau$ , while *deterministic* automata *never* contain  $\tau$  transitions. Moreover,  $P_\tau : \Sigma_\tau^* \rightarrow \Sigma^*$

is the *projection* that removes from strings in  $\Sigma_\tau^*$  all the  $\tau$  events.

When automata are brought together to interact, lock-step synchronization in the style of [15] is used.

*Definition 2:* Let  $G_1 = \langle \Sigma_1, Q_1, \rightarrow_1, Q_1^0, Q_1^m \rangle$  and  $G_2 = \langle \Sigma_2, Q_2, \rightarrow_2, Q_2^0, Q_2^m \rangle$  be two nondeterministic automata. The *synchronous composition* of  $G_1$  and  $G_2$  is defined as

$$G_1 \parallel G_2 := \langle \Sigma_1 \cup \Sigma_2, Q_1 \times Q_2, \rightarrow, Q_1^0 \times Q_2^0, Q_1^m \times Q_2^m \rangle \quad (1)$$

where

$$\begin{aligned} (x_1, x_2) &\xrightarrow{\sigma} (y_1, y_2) && \text{if } \sigma \in (\Sigma_1 \cap \Sigma_2), \\ &&& x_1 \xrightarrow{\sigma} y_1, \text{ and } x_2 \xrightarrow{\sigma} y_2; \\ (x_1, x_2) &\xrightarrow{\sigma} (y_1, x_2) && \text{if } \sigma \in (\Sigma_1 \setminus \Sigma_2) \cup \{\tau\} \\ &&& \text{and } x_1 \xrightarrow{\sigma} y_1; \\ (x_1, x_2) &\xrightarrow{\sigma} (x_1, y_2) && \text{if } \sigma \in (\Sigma_2 \setminus \Sigma_1) \cup \{\tau\} \\ &&& \text{and } x_2 \xrightarrow{\sigma} y_2. \end{aligned}$$

Importantly, synchronous composition only imposes lock-step synchronization on common events from  $\Sigma_1$  and  $\Sigma_2$ .

The transition relation of an automaton  $G$  is written in infix notation  $x \xrightarrow{\sigma} y$ , and it is extended to strings in  $\Sigma_\tau^*$  by letting  $x \xrightarrow{\varepsilon} x$  for all  $x \in Q$ , and  $x \xrightarrow{t} z$  if  $x \xrightarrow{t} y$  and  $y \xrightarrow{\sigma} z$  for some  $y \in Q$ . Furthermore,  $x \xrightarrow{t}$  means that  $x \xrightarrow{t} y$  for some  $y \in Q$ , and  $x \rightarrow y$  means that  $x \xrightarrow{t} y$  for some  $t \in \Sigma_\tau^*$ . These notations also apply to state sets, where  $X \xrightarrow{t} Y$  for  $X, Y \subseteq Q$  means that  $x \xrightarrow{t} y$  for some  $x \in X$  and  $y \in Y$ , and to automata, where  $G \xrightarrow{t}$  means that  $Q^0 \xrightarrow{t}$  ( $t$  is defined in  $G$ ) and  $G \xrightarrow{t} x$  means  $Q^0 \xrightarrow{t} x$ .

For brevity,  $p \xrightarrow{s} q$  for  $s \in \Sigma^*$  represents the existence of a string  $t \in \Sigma_\tau^*$  such that  $P_\tau(t) = s$  and  $p \xrightarrow{t} q$ . Thus,  $q \xrightarrow{u} p$  for  $u \in \Sigma_\tau^*$  means a path containing exactly the events in  $u$ , while  $q \xrightarrow{u} p$  for  $u \in \Sigma^*$  means existence of a path between  $p$  and  $q$  with an arbitrary number of  $\tau$  events between the observable events in  $u$ . Similarly,  $p \xrightarrow{\tau} q$  means the existence of a string  $t \in \{\tau\}^*$  such that  $p \xrightarrow{t} q$ .

The *language* of an automaton  $G$  is defined as  $\mathcal{L}(G) = \{s \in \Sigma^* \mid G \xrightarrow{s}\}$  and the language generated by  $G$  from  $q \in Q$  is  $\mathcal{L}(G, q) = \{s \in \Sigma^* \mid q \xrightarrow{s}\}$ , thus we do not include event  $\tau$  in the language of an automaton. Moreover, we also introduce projections  $P_i$  for  $i = 1, 2$ , which are  $P_i : (\Sigma_1 \cup \Sigma_2)^* \rightarrow \Sigma_i^*$  for  $i = 1, 2$ .

For a nondeterministic automaton  $G = \langle \Sigma_\tau, Q, \rightarrow, Q^0 \rangle$ , the set of *unobservably reached states* of  $B \subseteq 2^Q$ , is  $UR(B) = \bigcup \{C \subseteq Q \mid B \xrightarrow{\tau} C\}$ . Its *observer*  $det(G) = \langle \Sigma, X_{obs}, \rightarrow_{obs}, X_{obs}^0 \rangle$  is a deterministic automaton, where  $X_{obs}^0 = UR(Q^0)$  and  $X_{obs} \subseteq 2^Q$ , and  $X \xrightarrow{\sigma_{obs}} Y$ , where  $X, Y \in X_{obs}$ , if and only if  $Y = \bigcup \{UR(y) \mid x \xrightarrow{\sigma} y \text{ for some } x \in X \text{ and } y \in Q\}$ . By convention, only reachable states from  $X_{obs}^0$  under  $\rightarrow_{obs}$  are considered in this paper. We also refer to the observer as the (*current-state*) *estimator* of the system while an observer state is referred to as (*current-state*) *estimate*.

A common automaton operation is the *quotient* modulo, which is an equivalence relation on sets of states.

*Definition 3:* Let  $Z$  be a set. A relation  $\sim \subseteq Z \times Z$  is called an *equivalence relation* on  $Z$  if it is reflexive, symmetric, and transitive. Given an equivalence relation  $\sim$  on  $Z$ , the

equivalence class of  $z \in Z$  is  $[z] = \{z' \in Z \mid z \sim z'\}$ , and  $\bar{Z} = \{[z] \mid z \in Z\}$  is the set of all equivalence classes modulo  $\sim$ .

**Definition 4:** Let  $G = \langle \Sigma, Q, \rightarrow, Q^0 \rangle$  be an automaton and let  $\sim \subseteq Q \times Q$  be an equivalence relation. The *quotient automaton* of  $G$  modulo  $\sim$  is  $\tilde{G} = \langle \Sigma, \tilde{Q}, \rightarrow/\sim, \tilde{Q}^0 \rangle$ , where  $\rightarrow/\sim = \{([x], \sigma, [y]) \mid x \xrightarrow{\sigma} y\}$  and  $\tilde{Q}^0 = \{[x^0] \mid x^0 \in Q^0\}$ .

In order to compare automata structurally, we say that an automaton is a *subautomaton* of another automaton if all states and transitions in the first automaton are contained in the second one. Formally, we have the following definition:

**Definition 5:** Let  $G_1 = \langle \Sigma_\tau, Q_1, \rightarrow_1, Q_1^0, Q_1^m \rangle$  and  $G_2 = \langle \Sigma_\tau, Q_2, \rightarrow_2, Q_2^0, Q_2^m \rangle$  be two automata.  $G_1$  is a *subautomaton* of  $G_2$ , denoted by  $G_1 \sqsubseteq G_2$ , if  $Q_1 \subseteq Q_2$ ,  $\rightarrow_1 \subseteq \rightarrow_2$ ,  $Q_1^0 \subseteq Q_2^0$ , and  $Q_1^m \subseteq Q_2^m$ .

### B. Supervisory Control Theory

Considering plant  $G$  and specification  $K$ , *supervisory control theory* provides a method to synthesize a supervisor to restrict the behavior of the plant such that the given specification is always fulfilled. The supervisor  $S$  is a function defined from the language of the system  $G$  to the set of events, formally,  $S: \mathcal{L}(G) \rightarrow 2^\Sigma$ . We also partition the set of events as *uncontrollable events* and *controllable events*, i.e.,  $\Sigma = \Sigma_{uc} \cup \Sigma_c$ , where uncontrollable events cannot be disabled by the supervisor. In the figures the uncontrollable events are marked by an exclamation mark (!). The readers may refer to [4] for the main results of monolithic supervisory control under full observation. Here we focus on concepts and definitions relevant to the present paper and the synthesis procedure in this paper is done on deterministic automata. Two requirements for the supervisor are *controllability* and *nonblockingness*, where controllability captures *safety* in the presence of uncontrollable events and nonblockingness focuses on *liveness* of the system.

**Definition 6:** [4] Let  $G = \langle \Sigma, Q_G, \rightarrow_G, Q_G^0, Q_G^m \rangle$  and  $K = \langle \Sigma, Q_K, \rightarrow_K, Q_K^0, Q_K^m \rangle$  be two deterministic automata such that  $K \sqsubseteq G$ .  $K$  is *controllable* w.r.t.  $G$  if, for all states  $x \in Q_K$  and  $y \in Q_G$  and for every uncontrollable event  $v \in \Sigma_{uc}$  such that  $x \xrightarrow{v}_G y$ , it also holds that  $x \xrightarrow{v}_K y$ .

**Definition 7:** [4] Let  $G$  be a deterministic automaton. A state  $x$  is called *reachable* in  $G$  if  $G \rightarrow x$ , and *coreachable* in  $G$  if  $x \rightarrow Q^m$ . The automaton  $G$  is called *reachable* or *coreachable* if every state in  $G$  has this property.  $G$  is called *nonblocking* if every reachable state is coreachable.

The upper bound of controllable and nonblocking subautomata is again controllable and nonblocking. This implies the existence of a least restrictive subautomaton of the original system, which is achieved by the maximally-permissive and nonblocking supervisor.

**Definition 8:** Let  $G$  be an automaton, the supremal controllable and nonblocking subautomaton of  $G$  is called the *supremal supervisor*, denoted by  $\sup\mathcal{C}(G)$  where for all controllable and nonblocking automata  $K$  w.r.t.  $G$ ,  $K \sqsubseteq \sup\mathcal{C}(G)$ .

Synthesis of  $\sup\mathcal{C}(G)$  is done by iteratively removing blocking and uncontrollable states, until a fixed point is reached, and restricting the final automaton to the remaining states and their associated transitions, for more details please see [4], [13], [44].

In this paper, we assume that the modular system has a set of interacting components  $\{G_1, \dots, G_n\}$ , and there is also a set of supervisors in a modular structure, i.e.,  $\mathcal{S} = \{S_1, \dots, S_n\}$ . Here supervisor  $S_i$  is responsible for controlling  $G_i$ . The set of modular supervisors may be synchronized as  $\prod_{i=1}^n S_i$ .

### C. Opacity and Edit Functions

In this work, we suppose system  $G$  has certain secret information which is characterized by the set of states. Thus the state space is partitioned into two disjoint subsets:  $Q = Q^S \cup Q^{NS}$  where  $Q^S$  is the set of *secret states* capturing the secrets of the system, while  $Q^{NS}$  is the set of *non-secret states*. When the system  $\mathcal{G}$  is modular,  $\mathcal{G} = \{G_1, \dots, G_n\}$ , the set of secret states of the system,  $Q_S$ , is  $Q^S = \{(x_1, \dots, x_n) \mid \exists x_i \in Q_i^S\}$ .

Suppose there is an external intruder modeled as the observer of the system, which intends to infer the secrets of the system from its observation. Then a system is called *opaque* if the intruder is unable to determine unambiguously if the system has entered a secret state or not. Different notions of opacity have been introduced in literature and we focus on current-state opacity in this work.

**Definition 9:** A nondeterministic automaton  $G$  with a set of secret states  $Q^S$  is *current-state opaque* w.r.t.  $Q^S$  if  $(\forall s \in \mathcal{L}(G, q^0) : Q^0 \xrightarrow{s} Q^S)$  then  $[Q^0 \xrightarrow{s} Q^{NS}]$ .

The system is current-state opaque if for any string reaching a secret state there is string with the same sequence of observable events reaching a non-secret state. It is known that current-state opacity can be verified by building the standard observer automaton.

**Theorem 1:** Let  $G = \langle \Sigma_\tau, Q, \rightarrow, Q^0 \rangle$  be a nondeterministic automaton with set of secret states  $Q^S$ . Let  $\det(G) = \langle \Sigma, X_{obs}, \rightarrow_{obs}, X_{obs}^0 \rangle$  be the current-state estimator of  $G$ . Then  $G$  is current-state opaque w.r.t.  $Q^S$  if and only if  $[\det(G) \xrightarrow{s} X \text{ implies that } X \not\subseteq Q^S]$ .

If all states violating current-state opacity are removed from the observer  $\det(G)$ , then the accessible part of the remaining structure is called the *desired observer*, denoted by  $\det_d(G) = \langle \Sigma, X_{obsd}, \rightarrow_{obsd}, X_{obsd}^0 \rangle$ . The language generated by the desired observer is referred to as the *safe language*,  $L_{safe} = \mathcal{L}(\det_d(G))$ . Accordingly, we also define the *unsafe language*,  $L_{unsafe} = \mathcal{L}(G) \setminus L_{safe}$ .

If a system is not current state opaque then an interface based approach called *edit function* [20], [47] may be applied to enforce it. An edit function may insert events into the output of the system or erase events from the output of the system. It is assumed that the intruder fails to distinguish between an inserted event and its genuine counterpart. Let  $\Sigma' = \{\sigma \rightarrow \varepsilon : \sigma \in \Sigma\}$  be the set of “event erasure” events.

**Definition 10:** A *deterministic edit function* is defined as  $f_e : \Sigma^* \times \Sigma \rightarrow \Sigma^*$ . Given  $s \in \mathcal{L}(G)$ ,  $\sigma \in \Sigma$ ,

$$f_e(s, \sigma) = \begin{cases} s_I \sigma & \text{if } s_I \text{ is inserted before } \sigma \\ \varepsilon & \text{nothing is inserted and } \sigma \text{ is erased} \\ s_I & \text{if } s_I \text{ is inserted and } \sigma \text{ is erased} \end{cases}$$

With an abuse of notation, we also define a string-based edit function  $\hat{f}_e$  recursively as:  $\hat{f}_e(\varepsilon) = \varepsilon$ ,  $\hat{f}_e(s\sigma) = \hat{f}_e(s)f_e(s, \sigma)$

for  $s \in \Sigma^*$  and  $\sigma \in \Sigma$ . In the sequel, to ease the notational burden, we will drop the “ $\wedge$ ” in  $\hat{f}_e$  and it will be clear from the argument(s) of  $f_e$  which function we are referring to (incremental single-event one or string-based one).

Two notions termed *public safety* and *private safety* were defined in [20] to characterize the behavior of edit functions. In this paper, we consider private safety alone under the assumption that the intruder does not know about the implementation of an edit function.

**Definition 11 (Private Safety):** Given  $G$  and its observer  $\det(G)$ , an edit function  $f_e$  is privately safe if  $\forall s \in \mathcal{L}(\det(G))$ ,  $f_e(s) \in L_{\text{safe}}$ .

Recently a three-player game structure called *three-player observer (TPO)* w.r.t. the system was defined in [20] to embed edit functions. For the sake of completeness, we recall this definition (more details are available in [20]).

**Definition 12 (Three-player Observer):** Given a system  $G$  with its observer  $\det(G)$  and desired observer  $\det_d(G)$ , let  $I \subseteq X_{\text{obsd}} \times X_{\text{obs}}$  be the set of information states. A three-player observer w.r.t.  $G$  is a tuple of the form  $T = (Q_Y, Q_Z, Q_W, \Sigma, \Sigma', \Theta, \rightarrow_{yz}, \rightarrow_{zz}, \rightarrow_{zw}, \rightarrow_{wy}, y_0)$ , where:

- $Q_Y \subseteq I$  is the set of  $Y$  states.
- $Q_Z \subseteq I \times \Sigma$  is the set of  $Z$  states. Let  $I(z)$ ,  $E(z)$  denote the information state component and observable event component of a  $Z$  state  $z$  respectively, so that  $z = (I(z), E(z))$ .
- $Q_W \subseteq I \times (\Sigma \cup \Sigma')$  is the set of  $W$ -states. Let  $I(w)$ ,  $A(w)$  denote the information state component and action component of a  $W$  state  $w$  respectively, so that  $w = (I(w), A(w))$ .
- $\Sigma$  is the set of observable events.
- $\Sigma'$  is the set of event-erasure events.
- $\Theta \subseteq \Sigma \cup \{\varepsilon\} \cup \Sigma'$  is the set of edit decisions at  $Z$  states.
- (i)  $\rightarrow_{yz}: Q_Y \times \Sigma \times Q_Z$  is the transition function from  $Y$  states to  $Z$  states. For  $y = (x_d, x_f) \in Q_Y$ ,  $e_o \in \Sigma$ , we have:  $y \xrightarrow{e_o}_{yz} z \Rightarrow [x_f \xrightarrow{e_o}_{\text{obs}}] \wedge [I(z) = y] \wedge [E(z) = e_o]$ .
- (ii)  $\rightarrow_{zz}: Q_Z \times \Theta \times Q_Z$  is the transition function from  $Z$  states to  $Z$  states. For  $z = ((x_d, x_f), e_o) \in Q_Z$ ,  $\theta \in \Theta$ , we have:  $z \xrightarrow{\theta}_{zz} z' \Rightarrow [\theta \in \Sigma] \wedge [I(z') = (x'_d, x'_f)] \wedge [x_d \xrightarrow{\theta}_{\det_d} x'_d] \wedge [E(z') = e_o]$ .
- (iii)  $\rightarrow_{zw1}: Q_Z \times \Theta \times Q_W$  is the  $\varepsilon$  insertion transition function from  $Z$  states to  $W$  states. For  $z = ((x_d, x_f), e_o) \in Q_Z$ ,  $\theta \in \Theta$  we have:  $z \xrightarrow{\theta}_{zw1} w \Rightarrow [\theta = \varepsilon] \wedge [I(w) = I(z)] \wedge [A(w) = e_o] \wedge [x_d \xrightarrow{e_o}_{\det_d}] \wedge [x_f \xrightarrow{e_o}_{\text{obs}}]$ .
- (iv)  $\rightarrow_{zw2}: Q_Z \times \Theta \times Q_W$  is the event erasure transition function from  $Z$  states to  $W$  states. For  $z = ((x_d, x_f), e_o) \in Q_Z$ ,  $\theta \in \Theta$ , we have:  $z \xrightarrow{\theta}_{zw2} w \Rightarrow [\theta = e_o \rightarrow \varepsilon] \wedge [I(w) = I(z)] \wedge [A(w) = e_o \rightarrow \varepsilon] \wedge [x_f \xrightarrow{e_o}_{\text{obs}}]$ .
- (v)  $\rightarrow_{wy1}: Q_W \times \Sigma \times Q_Y$  is the transition function from  $W$  states whose action component is in  $\Sigma$  to  $Y$  states. For  $w = ((x_d, x_f), e_o) \in Q_W$ , we have:  $w \xrightarrow{e_o}_{wy1} y \Rightarrow [y = (x'_d, x'_f)] \wedge [x_d \xrightarrow{e_o}_{\det_d} x'_d] \wedge [x_f \xrightarrow{e_o}_{\text{obs}} x'_f]$ .
- (vi)  $\rightarrow_{wy2}: Q_W \times \Sigma \times Q_Y$  is the transition function from  $W$  states whose action component is in  $\Sigma'$  to  $Y$  states. For  $w = ((x_d, x_f), e_o \rightarrow \varepsilon) \in Q_W$ , we have:  $w \xrightarrow{e_o}_{wy2} y \Rightarrow [y = (x_d, x'_f)] \wedge [x_f \xrightarrow{e_o}_{\text{obs}} x'_f]$ .
- $y_0 = (x_{\text{obsd},0}, x_{\text{obs},0}) \in Q_Y$  is the initial state of  $T$ , where

$x_{\text{obsd},0}$  and  $x_{\text{obs},0}$  are initial states of  $\det_d(G)$  and  $\det(G)$ , respectively.

In general, a three-player observer characterizes a game between a dummy player, the edit function and the environment (system). The state space of a TPO is partitioned as:  $Q_Y$  states ( $Y$  states) where the dummy player plays;  $Q_Z$  states ( $Z$  states) where the edit function plays;  $Q_W$  states ( $W$  states) where the environment plays. A  $Y$  state contains the intruder's estimate (left component) as well as the system's true state estimate (right component). A  $\rightarrow_{yz}$  transition is defined out of a  $Y$  state, indicating that an observable event may occur and thus is received by the edit function. Then the TPO transits to a  $Z$  state and the turn of the game is passed to the edit function. Notice that the observable event does not really occur and this dummy player is only introduced to help determine the decisions of edit functions.

At a  $Z$  state, the edit function may choose to insert certain events (including  $\varepsilon$ ) or erase its last observed event. If a non- $\varepsilon$  event is inserted, a  $\rightarrow_{zz}$  transition leads the TPO to another  $Z$  state, which means the edit function still has the turn to insert more events until it decides to stop insertion by inserting  $\varepsilon$  or by erasing the last observed event. There may be multiple transitions defined out of a  $Z$  state, i.e., multiple edit decisions; we write  $\Theta(z)$  to denote the set of edit decisions defined at  $z \in Q_Z$  in a TPO.

If the edit function inserts nothing (respectively erases the event it receives from the dummy player), then a  $\rightarrow_{zw1}$  (respectively  $\rightarrow_{zw2}$ ) transition is defined and the TPO is at a  $W$  state. Then the environment plays by letting the observable event executed from its preceding  $Y$  state occur. Correspondingly, there are also two types of  $\rightarrow_{wy}$  transitions, where  $\rightarrow_{wy1}$  indicates that the executed observable event will be observed by the intruder while  $\rightarrow_{wy2}$  indicates that the executed observable event will not be observed by the intruder since it has been erased by the edit function.

When the three players take turns to play, the components of each player's states also get updated. From Def. 12, a  $\rightarrow_{yz}$  transition does not change the state estimates for the intruder or the system since the player at  $Y$  states is dummy and the observable events from  $Y$  states do not really occur. With a  $\rightarrow_{zz}$  transition, only  $x_d$  is updated since  $x_d$  is the estimate of the intruder and event insertion only alters the observation of the intruder. For  $\rightarrow_{zw}$  transitions, we only require the observable event to be defined at  $x_d$  or  $x_f$ . Finally, a  $\rightarrow_{wy1}$  transition updates both  $x_d$  and  $x_f$  while a  $\rightarrow_{wy2}$  transition only updates  $x_f$  as the intruder does not observe the erased event. To characterize the information flow in a TPO, the notion of *run* is defined in [20].

**Definition 13 (Run):** In a three-player observer  $T$ , a run is defined as:  $\omega = y_0 \xrightarrow{e_0} z_0^1 \xrightarrow{\theta_0^1} z_0^2 \xrightarrow{\theta_0^2} \dots \xrightarrow{\theta_0^{m_0-1}} z_0^{m_0} \xrightarrow{\theta_0^{m_0}} w_0 \xrightarrow{e_0} y_1 \xrightarrow{e_1} z_1^1 \xrightarrow{\theta_1^1} z_1^2 \xrightarrow{\theta_1^2} \dots \xrightarrow{\theta_1^{m_1-1}} z_1^{m_1} \xrightarrow{\theta_1^{m_1}} w_1 \xrightarrow{e_1} y_2 \dots \xrightarrow{e_n} z_n^1 \xrightarrow{\theta_n^1} \dots \xrightarrow{\theta_n^{m_n-1}} z_n^{m_n} \xrightarrow{\theta_n^{m_n}} w_n \xrightarrow{e_n} y_{n+1}$ , where  $y_0$  is the initial state of  $T$ ,  $e_i \in \Sigma$ ,  $\theta_i^j \in \Theta(z_i^j)$ ,  $\forall 0 \leq i \leq n$ ,  $1 \leq j \leq m_i$  and  $n \in \mathbb{N}$ ,  $m_i \in \mathbb{N}^+$ .

We let  $\Omega_T$  be the set of all runs in a TPO  $T$ . For simplicity, similar notations as for automata are defined for three-player observers and thus  $T \xrightarrow{\omega} x$  denotes the existence of a run in a

three-player observer. We also review the concepts of *string generated by a run* and *edit projection* defined in [20].

**Definition 14 (String Generated by a Run):** Given a run  $\omega$  as in Definition 13, the string generated by  $\omega$  is defined as:  $l(\omega) = \theta_0^1 \theta_0^2 \dots \theta_0^{m_0-1} \theta_0^{m_0} e_0 \theta_1^1 \dots \theta_1^{m_1} e_1 \dots e_{n-1} \theta_n^1 \dots \theta_n^{m_n} e_n$ , where  $\forall i \leq n$ ,  $\theta_i^{m_i} e_i = \varepsilon$  if  $\theta_i^{m_i} = e_i \rightarrow \varepsilon$ .

**Definition 15 (Edit Projection):** Given TPO  $T$  and run  $\omega_T$  as in Def. 13, the edit projection  $P_e : \Omega \rightarrow \mathcal{L}(G)$  is defined such that  $P_e(\omega_T) = e_0 e_1 \dots e_n$ .

In a TPO,  $y \in Q_Y$  is a *terminating state* if  $\nexists e_o \in \Sigma$ , s.t.  $y \xrightarrow{e_o}$ . And  $w \in Q_W$  is a *deadlocking state* if  $\nexists e_o \in \Sigma$ , s.t.  $w \xrightarrow{e_o} y$ . Also  $z \in Q_Z$  is a *deadlocking state* if  $\nexists \theta \in \Theta$ , s.t.  $z \xrightarrow{\theta} z'$  or  $z \xrightarrow{\theta} w$ . We call a TPO  $T$  *complete* [20] if there are no deadlocking  $W$  or  $Z$  states in  $T$  and  $\forall s \in \mathcal{L}(G)$ ,  $\exists \omega \in \Omega_T$ , s.t.  $P_e(\omega) = s$ .

**Definition 16 (Edit Function Embedded in TPO):** Given a TPO  $T$ , a deterministic edit function  $f_e$  is embedded in  $T$  if  $\forall s \in \mathcal{L}(G)$ ,  $\exists \omega \in \Omega_T$ , s.t.  $P_e(\omega) = s$  and  $l(\omega) = f_e(s)$ .

Next, we construct the *largest three-player observer* in the sense that all the other three-player observers are subautomata of it. Such a notion is well defined by considering *all* admissible transitions at *every* state of the TPO, according to the respective conditions in Def. 12.

Edit functions are designed to erase genuine events or insert fictitious ones to mislead the intruder. In theory, it is possible to design an edit function that erases all the events of the system, although this is not desirable. To avoid this situation usually the user provides some constraints on the edit functions. The constraint that is considered in this paper is to limit the number of consecutive erasures.

**Definition 17 (Edit Constraint):** The edit constraint, denoted by  $\Phi$ , requires that the edit function should not make  $n+1$  consecutive erasures where  $n \in \mathbb{N}$ .

Finally, we define the *All Edit Structure (AES)* [20] by considering the edit constraint. A synthesis procedure was also presented in [20] to construct the AES. Notice that the following definition is slightly different from the AES in the preliminary version of this work [26] since edit constraints are not considered in [26].

**Definition 18 (All Edit Structure):** Given system  $G$ , observer  $det(G)$  and desired estimator  $det_d(G)$ , the All Edit Structure is defined to be the largest complete three-player observer w.r.t.  $G$ , which satisfies the edit constraint.

From results in [20], private safety is achievable when the AES is not empty by construction. Hereafter, we assume that the AES is non-empty in the following discussion; if it is empty, then opacity cannot be enforced by the mechanism of edit functions. It was also proven in [20] that *all* privately safe edit functions satisfying edit constraints are embedded in the AES. Formally speaking, the following result holds.

**Theorem 2:** Given a system  $G$  and its corresponding AES under edit constraint  $\Phi$ , an edit function  $f_e$  is privately safe and satisfies  $\Phi$  if and only if  $f_e \in AES$ .

We end this section by briefly reviewing the pruning process discussed in [20] to construct the AES. The presence of edit constraints may preclude some undesired states from the AES, thus leaving some states without outgoing transitions, i.e., “deadlock”  $Z$  or  $W$  states. Those states reflect the inability of the edit function to issue a valid edit decision (for insertion

or erasure) while still maintaining opacity for all possible future behaviors, thus should be removed in the pruning process. Moreover,  $Y$  states that have transitions to a deadlock  $Z$  state need to be pruned as well, since  $Y$ -states are the states where the system issues an output event and the edit function is not allowed to prevent their occurrence.

The construction of the AES may also be interpreted as the calculation of a supervisor where the “plant” is the largest three-player observer in terms of subautomaton, including all potentially feasible edit decisions without considering edit constraints. The  $Y$  states are considered as marked states. The events labeling transitions from  $Y$  states to  $Z$  states and from  $W$  states to  $Y$  states are considered as uncontrollable, while the events labeling transitions from  $Z$  states to  $Z$  states and  $Z$  states to  $W$  states are viewed as controllable. We also define the proper specification by considering edit constraints, deleting states that violate them, and taking the trim of the resulting structure. The goal is to calculate the least restrictive, controllable and nonblocking supervisor based on the plant and this specification. Similar processes of pruning game structures akin to TPOs were discussed in prior work, e.g., [18], [20], [46]. We will leverage this approach in the following discussion, but in the framework of *modular supervisory control*.

### III. COMPOSITIONAL ABSTRACTION-BASED METHODOLOGY

This section presents our novel compositional and abstraction-based methodology for synthesizing modular form edit functions based on individual three-player observers after abstracting the original system. For simplicity, we call this methodology the CA-AES (Composition Abstraction-All Edit Structure) Algorithm hereafter. The input of the algorithm is a set of nondeterministic automata,  $\mathcal{G} = \{G_1, \dots, G_n\}$  and the output is a modular representation of edit functions, which is called *Modular Edit Structure*. The algorithm is summarized in Figure 1 and its steps are as follows. We will explain how to interpret the modular representation of edit functions later.

- (i) The algorithm first abstracts each individual automaton,  $G_i$ , using *opaque observation equivalence*. This results in  $\tilde{G}_i$ , which has fewer states and transitions compared to the original automaton.
- (ii) Next, we abstract the observer of  $\tilde{G}_i$ , i.e.,  $det(\tilde{G}_i)$ , by *opaque bisimulation* and *bisimulation*, resulting in two abstracted deterministic automata  $H_{i,ob}$  and  $H_{i,b}$ .
- (iii) Then we calculate the abstracted desired observer of  $G_i$  from  $H_{i,ob}$ , which is denoted by  $H_{i,obd}$ .
- (iv) Afterward, the largest (abstracted) three-player observer of each individual component  $G_i$  is calculated from the abstracted observer  $H_{i,b}$  and the abstracted desired observer  $H_{i,obd}$ , and it is denoted by  $TPO_i$ .
- (v) The final step is to calculate a modular nonblocking and controllable supervisor, then obtain a set of modular edit functions. This is done by transforming the largest three-player observers and the edit constraint to a set of automata, i.e.,  $G_i^T$  and  $K$ , respectively. This modular approach is in contrast to calculating monolithic edit functions embedded in the monolithic AES [20].

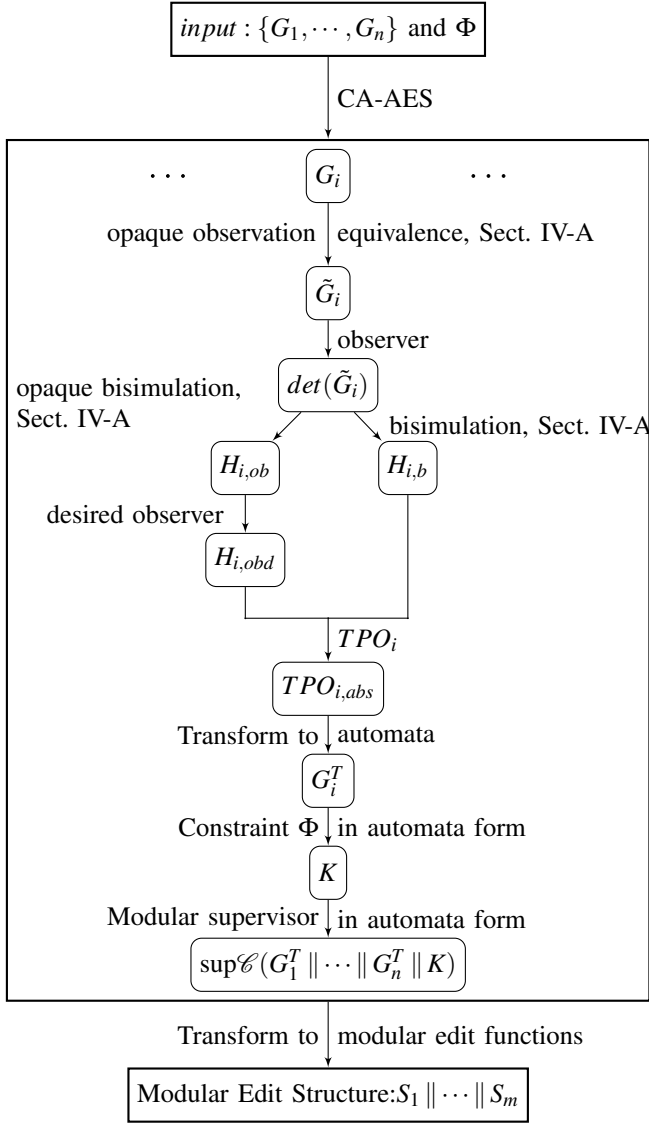


Fig. 1. The steps of Algorithm CA-AES.

Specifically, in step (iv), each abstracted three-player observer w.r.t. the corresponding individual system together with the constraint  $\Phi$  are transformed to a set of interacting automata. Then in step (v), a modular supremal controllable and nonblocking supervisor is calculated, thereby fulfilling the edit constraint in the composed structure. Consequently, the Modular Edit Structure is itself a modular supervisor. Regarding step (v), it is possible to leverage existing efficient algorithms on modular supervisory control to calculate the Modular Edit Structure.

In the monolithic approach of calculating the AES, individual systems  $G_1$  through  $G_n$  are synchronized first and then the observer of the synchronized system is built. Since the computational complexity of calculating the observer is exponential, synchronizing individual components before building the observer significantly increases the complexity, which may be  $2^{\prod_{i=1}^n |Q_i|}$  in the worst case where  $|Q_i|$  is the cardinality of  $Q_i$ . Moreover, constructing the All Edit Structure is polynomial in terms of the state space of the observer, which may be

potentially intractable when we deal with the synchronized system. In contrast, our compositional and abstraction-based approach reduces computational cost considerably both from abstracting individual systems and conducting computation in a modular way. However, as will be demonstrated later, some edit decisions may be omitted in the Modular Edit Structure output by Algorithm CA-AES.

The presented approach relies heavily on the use of three-player observers. We present an example to better understand the structure of such observers.

*Example 1:* Consider the nondeterministic automaton  $G_1$  with secret states set  $Q_1^S = \{q_3\}$ , shown in Fig. 2. To generate the three-player observer of  $G_1$ , first the observer of  $G_1$  needs to be built, which is shown as  $det(G_1)$  in Fig. 2. To generate the desired observer the state  $\{q_3\} \subseteq Q_1^S$  needs to be removed. The desired observer  $det_d(G_1)$  is shown in Fig. 2.

Then we follow the procedures in [20] to build the TPO w.r.t.  $det(G_1)$  in Fig. 2 (labeled as  $T_1'$ ). As is discussed, the game on the TPO is initiated from  $Y$ -state  $(q_0, q_0)$  where the dummy player executes the observable event  $\gamma$  (the only event defined at  $q_0$  in  $det(G_1)$ ). Then the edit function takes the turn to play at the  $Z$  state  $(q_0, q_0, \gamma)$  where it has two choices: insert nothing or erase  $\gamma$ . If  $\gamma$  is erased, then the  $W$  state  $(q_0, q_0, \gamma \rightarrow \epsilon)$  is reached where the environment plays by executing  $\gamma$ . Then the turn is passed back to the dummy player and the rest of the structure is interpreted similarly.

The compositional abstraction-based approach is explained in more details in the following sections. First, in Section IV, we discuss abstractions at the component level and synchronization of individual three-player observers, formalizing steps (i)-(iv) of Algorithm CA-AES. Then, in Section V, we discuss the last step of Algorithm CA-AES.

#### IV. SYNCHRONIZATION AND ABSTRACTION OPERATIONS

This section presents results on abstraction and composition that support steps (i)-(iv) of Algorithm CA-CAS. First, Sect. IV-A describes the methods to abstract nondeterministic automata and their observers. Next, Sect. IV-B describes the process of transforming every individual three-player observer to an automaton form and shows that the automaton representation is a substructure (in the sense of subgraph) of the largest monolithic three-player observer.

##### A. Opaque observation equivalence

The first strategy used in Algorithm CA-AES to alleviate state space explosion is abstraction of system components. This subsection contains a collection of abstraction methods that can be used to abstract nondeterministic automata and their observers such that the abstracted observers and the desired observers are bisimilar to their original counterparts. The abstraction methods are based on bisimulation and observation equivalence, which are computationally efficient and can be calculated in polynomial-time [12]. We will prove in Theorem 5 that if we build the largest three-player observer based on the abstracted observer and the desired observer, we obtain the same runs, consequently the same edit functions

as we do from the largest three-player observer based on the original observer and desired observer.

Bisimulation is a widely-used notion of abstraction that merges states with the same future behavior.

**Definition 19:** [25] Let  $G = \langle \Sigma_\tau, Q, \rightarrow, Q^0 \rangle$  be a non-deterministic automaton. An equivalence relation  $\approx \subseteq Q \times Q$  is called a *bisimulation* on  $G$ , if the following holds for all  $x_1, x_2 \in Q$  such that  $x_1 \approx x_2$ : if  $x_1 \xrightarrow{\sigma} y_1$  for some  $\sigma \in \Sigma_\tau$ , then there exists  $y_2 \in Q$  such that  $x_2 \xrightarrow{\sigma} y_2$ , and  $y_1 \approx y_2$ .

Bisimulation seeks to merge states with the same outgoing transitions to equivalent states *including unobservable events*, i.e.,  $\tau$  events. If the unobservable events are disregarded, a more general abstraction method called *weak bisimulation* or *observation equivalence* naturally comes [25].

**Definition 20:** Let  $G = \langle \Sigma_\tau, Q, \rightarrow, Q^0 \rangle$  be a nondeterministic automaton. An equivalence relation  $\sim \subseteq Q \times Q$  is called an *observation equivalence* on  $G$ , if the following holds for all  $x_1, x_2 \in Q$  such that  $x_1 \sim x_2$ : if  $x_1 \xrightarrow{s} y_1$  for some  $s \in \Sigma^*$ , then there exists  $y_2 \in Q$  such that  $x_2 \xrightarrow{s} y_2$ , and  $y_1 \sim y_2$ .

In order to use observation equivalence for abstraction in the opacity setting, the set of secret states needs to be taken into account. In the following discussion, a restricted version of observation equivalence called *opaque observation equivalence* is employed. This notion was first defined in [27] in the context of verifying opacity.

**Definition 21:** Let  $G = \langle \Sigma_\tau, Q, \rightarrow, Q^0 \rangle$  be a nondeterministic automaton with set of secret states  $Q^S \subseteq Q$  and set of non-secret states  $Q^{NS} = Q \setminus Q^S$ . An equivalence relation  $\sim_o \subseteq Q \times Q$  is called an *opaque observation equivalence* on  $G$  with respects to  $Q^S$ , if the following holds for all  $x_1, x_2 \in Q$  such that  $x_1 \sim_o x_2$ :

- (i) if  $x_1 \xrightarrow{s} y_1$  for some  $s \in \Sigma^*$ , then there exists  $y_2 \in Q$  such that  $x_2 \xrightarrow{s} y_2$ , and  $y_1 \sim_o y_2$ ,
- (ii)  $x_1 \in Q^S$  if and only if  $x_2 \in Q^S$ .

We also wish to use bisimulation to abstract the observer of a nondeterministic system. Besides opaque observation equivalence, *opaque bisimulation* is also defined.

**Definition 22:** Let  $G = \langle \Sigma_\tau, Q, \rightarrow, Q^0 \rangle$  be a nondeterministic automaton with set of secret states  $Q^S \subseteq Q$  and set of non-secret states  $Q^{NS} = Q \setminus Q^S$ . Let  $det(G) = \langle \Sigma, X_{obs}, \rightarrow_{obs}, X_{obs}^0 \rangle$  be the observer of  $G$ . An equivalence relation  $\approx_o \subseteq X_{obs} \times X_{obs}$  is called an *opaque bisimulation equivalence* on  $det(G)$  with respects to  $Q^S$ , if the following holds for all  $X_1, X_2 \in X_{obs}$  such that  $X_1 \approx_o X_2$ :

- (i) if  $X_1 \xrightarrow{s} Y_1$  for some  $s \in \Sigma^*$ , then there exists  $Y_2 \in X_{obs}$  such that  $X_2 \xrightarrow{s} Y_2$ , and  $Y_1 \approx_o Y_2$ ,
- (ii)  $X_1 \subseteq Q^S$  if and only if  $X_2 \subseteq Q^S$ .

The first step of Algorithm CA-AES is to abstract the system using opaque observation equivalence. It has been shown in [32] that if two automata are bisimilar, then their observers are also bisimilar. In this paper this result is extended such that abstracting a nondeterministic automaton using opaque observation equivalence results in an observer and a desired observer which are bisimilar to the observer and the desired observer of the original system, respectively.

**Proposition 3:** Let  $G = \langle \Sigma_\tau, Q, \rightarrow, Q^0 \rangle$  be a nondeterministic automaton with set of secret states  $Q^S \subseteq Q$  and set of non-secret states  $Q^{NS} = Q \setminus Q^S$ . Let  $\sim_o$  be an opaque observation

equivalence on  $G$  resulting in  $\tilde{G}$  and let  $\approx$  be a bisimulation. Let  $det_d(G)$  and  $det_d(\tilde{G})$  be the desired observer of  $G$  and  $\tilde{G}$ . Then  $det(G) \approx det(\tilde{G})$  and  $det_d(G) \approx det_d(\tilde{G})$ .

*Proof:* First we prove that  $det(G) \approx det(\tilde{G})$ . To prove  $det(G) \approx det(\tilde{G})$  it is enough to show that  $det(G) \xrightarrow{s} X$  if and only if  $det(\tilde{G}) \xrightarrow{s} \tilde{X}$ , which implies language equivalence between  $det(G)$  and  $det(\tilde{G})$  since  $det(G)$  and  $det(\tilde{G})$  are deterministic. This can be shown by induction. Moreover, in the induction we also show that  $x \in X$  if and only if there exist  $[x'] \in \tilde{X}$  such that  $x \in [x']$ . This is used for the second part of the proof, where we show  $det_d(G) \approx det_d(\tilde{G})$ .

It is shown by induction on  $n \geq 0$  that  $X^0 \xrightarrow{\sigma_1} X^1 \xrightarrow{\sigma_2} \dots \xrightarrow{\sigma_n} X^n$  in  $det(G)$  if and only if  $\tilde{X}^0 \xrightarrow{\sigma_1} \tilde{X}^1 \xrightarrow{\sigma_2} \dots \xrightarrow{\sigma_n} \tilde{X}^n$  in  $det(\tilde{G})$  such that  $x \in X^j$  if and only if  $[x'] \in \tilde{X}^j$ , where  $x \in [x']$ , for  $1 \leq j \leq n$ .

*Base case:*  $n = 0$ . Let  $X^0$  be the initial state of  $det(G)$  and  $\tilde{X}^0$  be the initial state of  $det(\tilde{G})$ . It is shown that  $x \in X^0$  if and only if there exists  $[x'] \in \tilde{X}^0$  such that  $x \in [x']$ .

First, let  $x \in X^0$ . Then based on  $UR(x^0)$ , it follows that there exists  $x^0 \in Q^0$  such that  $x^0 \xrightarrow{s} x$  in  $G$ . Since  $G \sim_o \tilde{G}$  then based on Def. 21, there exists  $[x^0] \in \tilde{X}^0$  such that  $[x^0] \xrightarrow{s} [x']$  in  $\tilde{G}$  such that  $x^0 \in [x^0]$  and  $x \in [x']$ . Then based on  $UR(x^0)$  it follows that  $[x'] \in \tilde{X}^0$ .

Now let  $[x'] \in \tilde{X}^0$ . Then based on  $UR(x^0)$ , it follows that there exists  $[x^0] \in Q^0$  such that  $[x^0] \xrightarrow{s} [x']$  in  $\tilde{G}$ . Since  $G \sim_o \tilde{G}$  then based on Def. 21, there exists  $x^0 \in X^0$  such that  $x^0 \xrightarrow{s} x$  in  $G$  such that  $x^0 \in [x^0]$  and  $x \in [x']$ . Then based on  $UR(x^0)$  it follows that  $x \in X^0$ .

*Inductive step:* Assume the claim holds for some  $n \geq 0$ , i.e.,  $X^0 \xrightarrow{\sigma_1 \sigma_2 \dots \sigma_n} X^n = X$  in  $det(G)$  if and only if  $\tilde{X} = \tilde{X}^0 \xrightarrow{\sigma_1 \sigma_2 \dots \sigma_n} \tilde{X}^n = \tilde{X}$  in  $det(\tilde{G})$ , such that  $x \in X^k$  if and only if there exists  $[x'] \in \tilde{X}^k$  such that  $x \in [x']$  for all  $0 \leq k < n^1$ . It must be shown that  $X = X^n \xrightarrow{\sigma_{n+1}} Y$  in  $det(G)$  if and only if  $\tilde{X} = \tilde{X}^n \xrightarrow{\sigma_{n+1}} \tilde{Y}$  in  $det(\tilde{G})$  such that  $x \in X$  if and only if there exists  $[x'] \in \tilde{X}$  such that  $x \in [x']$ .

First, let  $X = X^n \xrightarrow{\sigma_{n+1}} Y$  in  $det(G)$  and let  $x \in X$ . Then based on  $UR(x)$  it holds that  $x = x^1 \xrightarrow{s} \dots \xrightarrow{s} x^r \xrightarrow{\sigma_{n+1}} y$  in  $G$ , where  $x^j \in X$  for all  $1 \leq j \leq r$  and  $y \in Y$ . Since  $G \sim_o \tilde{G}$  it holds that  $[x^j] = [x^1] \xrightarrow{s} \dots \xrightarrow{s} [x^r] \xrightarrow{\sigma_{n+1}} [y']$  in  $\tilde{G}$  such that  $x^j \in [x^j]$  for all  $1 \leq j \leq r$  and  $y \in [y']$ . Based on  $UR(x)$  and inductive assumption it holds that  $det(\tilde{G}) \xrightarrow{\sigma_1 \sigma_2 \dots \sigma_n} \tilde{X}^n = \tilde{X} \xrightarrow{\sigma_{n+1}} \tilde{Y}$  and  $[x'] \in \tilde{X}$ .

Now let  $\tilde{X} = \tilde{X}^n \xrightarrow{\sigma_{n+1}} \tilde{Y}$  in  $det(\tilde{G})$  and let  $[x] \in \tilde{X}$ . Then based on  $UR(x)$  it holds  $[x] = [x^1] \xrightarrow{s} \dots \xrightarrow{s} [x^r] \xrightarrow{\sigma_{n+1}} [y]$  in  $\tilde{G}$ , where  $[x^i] \in \tilde{X}$  for all  $1 \leq i \leq r$  and  $[y] \in \tilde{Y}$ . Since  $G \sim_o \tilde{G}$  it holds that  $x' = x^1 \xrightarrow{s} \dots \xrightarrow{s} x^r \xrightarrow{\sigma_{n+1}} y'$  in  $G$  such that  $x^i \in [x^i]$  for all  $1 \leq i \leq r$  and  $y' \in [y]$ . Based on  $UR(x)$  and inductive assumption it holds that  $det(G) \xrightarrow{\sigma_1 \sigma_2 \dots \sigma_n} X^n = X \xrightarrow{\sigma_{n+1}} Y$  such that  $x' \in X$ .

Now we need to show that  $det_d(G) \approx det_d(\tilde{G})$ . It was proven above that  $det(G) \approx det(\tilde{G})$ , which means  $det(G) \xrightarrow{s} X$  if and only if  $det(\tilde{G}) \xrightarrow{s} \tilde{X}$  and  $x \in X$  if and only if  $[x'] \in \tilde{X}$ , where  $x \in [x']$ . Therefore, it is enough to show that  $X \notin X_{obsd}$  if and only if  $\tilde{X} \notin \tilde{X}_{obsd}$ .

<sup>1</sup>Since the base case of the induction is proven for  $n = 0$ ,  $X^0 \xrightarrow{s} X$ , the inductive step is considered true for  $0 \leq k < n$ .

First assume  $X \subseteq Q^S$ , which means for all  $x \in X$  it holds that  $x \in Q^S$  and  $X \notin X_{obsd}$ . Since for all  $x \in X$  it holds that there exist  $[x'] \in \tilde{X}$  such that  $x \in [x']$  then based on Def. 21 it holds that  $[x'] \in \tilde{Q}^S$ . Thus, it can be concluded that for all  $[x'] \in \tilde{X}$  it holds that  $[x'] \in \tilde{Q}^S$ . This means that  $\tilde{X} \subseteq \tilde{Q}^S$  and consequently  $\tilde{X} \notin \tilde{X}_{obsd}$ .

Now assume  $\tilde{X} \subseteq \tilde{Q}^S$ , which means for all  $[x'] \in \tilde{X}$  it holds that  $[x'] \in \tilde{Q}^S$  and  $\tilde{X} \notin \tilde{X}_{obsd}$ . If  $[x'] \in \tilde{Q}^S$  then for all  $x \in [x']$  it holds that  $x \in Q^S$ . Moreover, it was shown above that  $[x'] \in \tilde{X}$  if and only if  $x \in X$ , where  $x \in [x']$ . Thus, from  $\tilde{X} \subseteq \tilde{Q}^S$  it follows that  $X \subseteq Q^S$ , which means that  $X \notin X_{obsd}$ .

Thus, it can be concluded that  $det_d(G) \approx det_d(\tilde{G})$ . ■

Opaque observation equivalence seeks to merge states of a nondeterministic automaton, which are “equivalent”, before constructing the observer. After calculating the observer, it is possible to further abstract the observer using opaque bisimulation. This guarantees that the smallest abstracted observer generates the same language as the original observer. In the following, Proposition 4 shows that if opaque bisimulation is used to abstract the observer, then the abstracted desired observer is also bisimilar to the original desired observer.

**Proposition 4:** Let  $G = \langle \Sigma, \tau, Q, \rightarrow, Q^0 \rangle$  be a nondeterministic automaton with set of secret states  $Q^S \subseteq Q$  and set of non-secret states  $Q^{NS} = Q \setminus Q^S$ . Let  $\approx_o$  be an opaque bisimulation on  $det(G)$  resulting in  $\widetilde{det(G)}$ . Let  $det_d(G)$  and  $H_d$  be the desired observers of  $det(G)$  and  $\widetilde{det(G)}$ , respectively. Then  $det_d(G) \approx H_d$ , where  $\approx$  is a bisimulation relation.

*Proof:* Since  $det(G) \approx_o \widetilde{det(G)}$  based on Def. 22 it holds that  $det(G) \xrightarrow{s} X$  if and only if  $\widetilde{det(G)} \xrightarrow{s} [X']$  and  $X \in [X']$ . Thus, it is enough show that  $X \notin X_{obs, det_d(G)}$  if and only if  $[X'] \notin X_{obs, H_d}$ , where  $X \in [X']$ .

First assume  $X \subseteq Q^S$ , so  $X \notin X_{obs, det_d(G)}$ . Then since  $X \in [X']$  based on Def. 22 it holds that for all  $X' \in [X']$ ,  $X' \subseteq Q^S$ . This means  $[X'] \subseteq Q^S$  and consequently  $[X'] \notin X_{obs, H_d}$ .

Then assume  $[X'] \subseteq Q^S$ , so  $[X'] \notin X_{obs, H_d}$ . Since  $X \in [X']$  based on Def. 22,  $X \subseteq Q^S$  holds, i.e.,  $X \notin X_{obs, det_d(G)}$ . ■

We now present the main results of this subsection.

**Theorem 5:** Let  $G$  be a nondeterministic automaton with secret states  $Q^S \subseteq Q$  and non-secret states  $Q^{NS} = Q \setminus Q^S$ . Let  $det(G)$  and  $det_d(G)$  be the observer and the desired observer of  $G$ , respectively. Let  $\sim_o$  be an opaque observation equivalence on  $G$  such that  $\tilde{G} \sim_o G$ . Let  $H_{ob} \approx_o det(\tilde{G})$  and  $H_b \approx det(\tilde{G})$  where  $\approx_o$  and  $\approx$  are opaque bisimulation and bisimulation, respectively. Let  $H_{obd}$  be the desired observer of  $H_{ob}$ . Let  $T$  be the largest three-player observer w.r.t.  $det(G)$  and  $det_d(G)$ , also let  $T'$  be the largest three-player observer w.r.t.  $H_{obd}$  and  $H_b$ . Then  $T \xrightarrow{o} q$  if and only if  $T' \xrightarrow{o} \tilde{q}$ .

*Proof:* In Supplementary Materials. ■

Theorem 5 proves that the largest three-player observer obtained from the abstracted system (using opaque observation equivalence and opaque bisimulation) has the same set of runs with that obtained from the original system. This result is essential for the correctness of Algorithm CA-AES.

**Remark 1:** The abstractions in the worst case scenario fail to merge any states. However, as pointed out in the paper the complexity of the abstraction methods is polynomial, while the complexity of calculating the observer is exponential in the

number of states. Thus, if the abstraction results in merging even few states, it can potentially reduce the complexity of calculating the observer significantly. Therefore, it is worth applying the abstraction algorithm before calculating the observers.

**Example 2:** Consider the nondeterministic system  $\mathcal{G} = \{G_1, G_2\}$ , shown in Fig. 2, with secret states sets  $Q_1^S = \{q_3\}$  and  $Q_2^S = \{s_3\}$  where all the events are observable except event  $\tau$ . In  $G_1$  states  $q_1$  and  $q_2$  are opaque observation equivalent as they both have the same secrecy status and equivalent states can be reached from both,  $q_1 \xrightarrow{\alpha} q_3$  and  $q_2 \xrightarrow{\alpha} q_3$ , and  $q_1 \xrightarrow{\tau} q_2$  and  $q_2 \xrightarrow{\tau} q_1$ . Merging  $q_1$  and  $q_2$  results in the abstracted automaton  $\tilde{G}_1$  shown in Fig. 2. Moreover, states  $s_1$  and  $s_2$  are also opaque observation equivalent and merging them results in automaton  $\tilde{G}_2$  shown Fig. 2. After abstracting the automata, the system becomes a deterministic system. Moreover, the observers as of  $\tilde{G}_1$  and  $\tilde{G}_2$  are bisimilar to  $det(G_1)$  and  $det(G_2)$ , respectively. The same is also true for the desired observer of  $\tilde{G}_1$  and  $\tilde{G}_2$ . Fig. 2 shows the largest three-player observers of  $\tilde{G}_1$  and  $\tilde{G}_2$ , respectively.

## B. Synchronous composition of TPOs

The second strategy used in Algorithm CA-AES to reduce computation complexity is synchronous composition of individual systems. In this work, the main advantage of our compositional approach is to build the largest three-player observer of each component individually, instead of synchronizing individual components and then building the largest monolithic three-player observer. Before synchronization, we first transfer each individual TPO to an automaton using Def. 23. Next, the individual automata are transformed to a set of interacting automata based on Def. 24. It is shown in Theorem 8 that the set of modular three-player observers form a subsystem of their monolithic counterpart, in the sense that some runs are omitted after synchronization. Before Theorem 8, Lemmas 6 and 7 establish that synchronization of individual observers (respectively, desired observers) is isomorphic to the observer (respectively, desired observers) of the synchronized system.

**Definition 23:** Let  $T = \langle Q_Y, Q_Z, Q_W, \Sigma, \Sigma^E, \Theta, \rightarrow_{yz}, \rightarrow_{zz}, \rightarrow_{zw}, \rightarrow_{wy}, y_0 \rangle$  be a three-player observer. Automaton  $M^T = \langle \Sigma_{M^T}, Q, \rightarrow, Q^0, Q^m \rangle$  is the monolithic transformed deterministic automaton of  $T$  where,

- (i)  $\Sigma_{M^T} = \Sigma \cup [ \bigcup_{p \in Q_Z} \Theta_{E(p)} ] \cup [ \bigcup_{p \in Q_W} \Sigma_{A(p), w} ]$
- (ii)  $Q = Q_Y \cup Q_Z \cup Q_W$ ,
- (iii)  $\rightarrow = \{ (p, \alpha, q) \mid p \in Q_Y \wedge p \xrightarrow{\alpha}_{yz} q \} \cup \{ (p, \sigma, q) \mid p \in Q_Z \wedge \sigma = \alpha_{E(p)} \wedge p \xrightarrow{\alpha}_{zz, zw_1, zw_2} q \} \cup \{ (p, \sigma, q) \mid p \in Q_W \wedge \sigma = \alpha_{A(p), w} \wedge p \xrightarrow{\alpha}_{wy_1, wy_2} q \}$
- (iv)  $Q^0 = y_0$ ,
- (v)  $Q^m = Q_Y$ .

The events labeling outgoing transitions mapped from original  $Y$  states in  $T$ , i.e.,  $\{ (p, \alpha, q) \mid p \in Q_Y \wedge p \xrightarrow{\alpha}_{yz} q \}$  and outgoing transitions mapped from original  $W$  states in  $T$ ,  $\{ (p, \sigma, q) \mid p \in Q_W \wedge \sigma = \alpha_{A(p), w} \wedge p \xrightarrow{\alpha}_{wy_1, wy_2} q \}$  are considered as uncontrollable while the other events are controllable.

In Def. 23,  $\Sigma_\alpha$  represents that  $\alpha$  is added to all the events of  $\Sigma$ . To transform a three-player observer to an automaton,



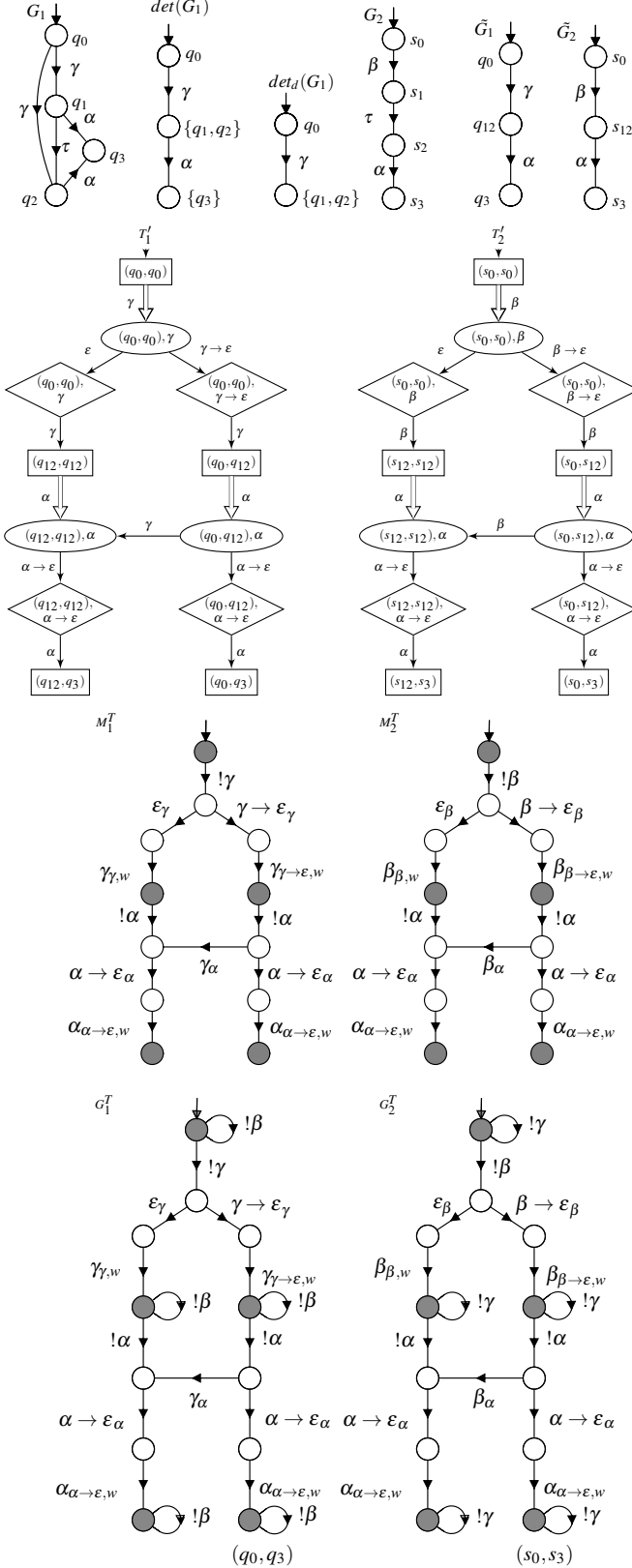


Fig. 2. System  $\mathcal{G} = \{G_1, G_2\}$  and its abstraction  $\{\tilde{G}_1, \tilde{G}_2\}$ . The figure also shows the largest three-player observers  $T_1'$  and  $T_2'$  of the abstracted components and their automata transformations, denoted by  $G_1^T$  and  $G_2^T$ . The uncontrollable events are marked by (!).

each state of the three-player observer is considered as an automaton state,  $Q_i = Q_Y^i \cup Q_Z^i \cup Q_W^i$  in Def. 23. Moreover, the information about the states needs to be considered to distinguish some transitions in the transformed automaton and to have a correct synchronization of three-player observers, since the information about state types ( $Y, W, Z$ ) is lost in the transformation. To this end, in the transformed automaton, the events labeling the transitions from  $z$  to  $z$  states, from  $z$  to  $w$  states and from  $w$  to  $y$  states, need to have the information of the predecessor states reflected in them. Thus, the events in the transformed automaton have the observable event components of  $Z$  states,  $\bigcup_{p \in Q_Z^i} \Theta_{E(p)}^i$ , and the action components of  $W$  states,

$\bigcup_{p \in Q_Z^i} \Sigma_{A(p), w}^i$ . The initial state of the transformed automaton  $p \in Q_Z^i$  is the initial state of the three-player observer and the marked states are the original  $Y$  states.

**Example 3:** Consider the two three-player observers  $T_1'$  and  $T_2'$  shown in Fig. 2. To transform the three-player observers  $T_1$  and  $T_2$  to their monolithic automata, renaming  $\rho : \{\alpha, \beta, \gamma, \varepsilon_\gamma, \gamma \rightarrow \varepsilon_\gamma, \gamma \rightarrow \varepsilon_{\gamma, w}, \gamma \varepsilon, w, \gamma \alpha, \varepsilon_\beta, \beta \rightarrow \varepsilon_\beta, \beta \rightarrow \varepsilon_{\beta, w}, \beta \varepsilon, w, \beta \alpha, \alpha \rightarrow \varepsilon_\alpha, \alpha \rightarrow \varepsilon_{\alpha, w}\} \rightarrow \{\alpha, \beta, \gamma, \varepsilon, \alpha \rightarrow \varepsilon, \beta \rightarrow \varepsilon, \gamma \rightarrow \varepsilon\}$  is introduced. Next, all  $Y$  states in  $T_1'$  and  $T_2'$  are considered as marked. Fig. 2 shows  $M_1^T$  and  $M_2^T$ , which are the monolithic transformed automata from  $T_1'$  and  $T_2'$ , respectively. All the events that come from transitions defined out of  $Y$  states and  $W$  states in  $T_1'$  and  $T_2'$  are considered as uncontrollable in  $G_1^T$  and  $G_2^T$ , thus  $\alpha, \beta$  and  $\gamma$  are uncontrollable.

This paper describes the compositional approach for modular systems. In order to have a correct interaction between the transformed automata, the transformation of the three-player observers needs to be done in the modular setting.

**Definition 24:** Let  $\mathcal{T} = \{T_1, \dots, T_n\}$  be a three-player observer system such that  $T_i = \langle Q_Y^i, Q_Z^i, Q_W^i, \Sigma^i, \Sigma_i^\varepsilon, \Theta^i, \rightarrow_{yz}^i, \rightarrow_{zz}^i, \rightarrow_{zw}^i, \rightarrow_{wy}^i, y_0^i \rangle$ . Let  $M_i^T = \langle \Sigma_{MT}^i, Q_M^i, \rightarrow_M^i, Q_M^0, Q_M^m \rangle$  be the monolithic transformed deterministic automaton of  $T_i$ , based on Def. 23. The transformed automaton system of  $\mathcal{T}$  is  $\mathcal{G}^T = \{G_1^T, \dots, G_n^T\}$ , where  $G_i^T = \langle \Sigma_{GT}^i, Q_i, \rightarrow^i, Q_i^0, Q_i^m \rangle$  and

- (i)  $\Sigma_{GT}^i = \Sigma_{MT}^i \cup \left[ \bigcup_{\alpha \in (\Sigma^j \setminus \Sigma^i); j \neq i} ((\Sigma^i \cap \Sigma^j)_\alpha \cup (\Sigma^i \cap \Sigma^j)_{\alpha, w} \cup (\Sigma^i \cap \Sigma^j)_{\alpha \rightarrow \varepsilon, w}) \right]$
- (ii)  $Q_i = Q_M^i$ ,
- (iii)  $\rightarrow_M^i \cup \{(p, \alpha, p) \mid p \in Q_Y^i \text{ and } \alpha \in \bigcup_{j \neq i} (\Sigma^j \setminus \Sigma^i)\}$
- (iv)  $Q_i^0 = Q_M^0$ ,
- (v)  $Q_i^m = Q_M^m$ .

Since some shared events in the transformed automaton become local after incorporating the extra state information, they need to be added in the alphabet of the transformed automaton,  $\bigcup_{\alpha \in (\Sigma^j \setminus \Sigma^i); j \neq i} ((\Sigma^i \cap \Sigma^j)_\alpha \cup (\Sigma^i \cap \Sigma^j)_{\alpha, w} \cup (\Sigma^i \cap \Sigma^j)_{\alpha \rightarrow \varepsilon, w})$  in Def. 24. Moreover, the events not defined from  $Y$  states of certain TPO  $T$  but defined from  $Y$  states of some other TPO  $T'$  are added as self-loops at the corresponding states in the transformed automaton of  $T$ ,  $\{(p, \alpha, p) \mid p \in Q_Y^i \text{ and } \alpha \in \bigcup_{j \neq i} (\Sigma^j \setminus \Sigma^i)\}$ . To create a map between the events of a TPO and its transformed automaton, renaming of events is necessary. Note, when the transformation of a single automaton

TABLE I

THE TABLE SHOWS THE LINK BETWEEN THE EVENTS OF A TPO, ITS TRANSFORMED AUTOMATON AND THE CORRESPONDING RENAMING.

TPO $T$	Automaton $G^T$	Renaming $\rho$
$y \xrightarrow{\alpha} yz z$	$y \xrightarrow{\alpha} z$	$\rho(\alpha) = \alpha$
$z \xrightarrow{\alpha} z z z', E(z) = e_o$	$z \xrightarrow{\alpha_{e_o}} z'$	$\rho(\alpha_{e_o}) = \alpha$
$z \xrightarrow{\alpha} z w_1 w, E(z) = e_o$	$z \xrightarrow{\alpha_{e_o}} w$	$\rho(\alpha_{e_o}) = \alpha$
$z \xrightarrow{\alpha} z w_2 w, E(z) = e_o$	$z \xrightarrow{\alpha_{e_o}} w$	$\rho(\alpha_{e_o}) = \alpha$
$w \xrightarrow{\alpha} w y_1 y, A(w) = e_o$	$w \xrightarrow{\alpha_{e_o, w}} y$	$\rho(\alpha_{e_o, w}) = \alpha$
$w \xrightarrow{\alpha} w y_2 y, A(w) = e_o \rightarrow \varepsilon$	$w \xrightarrow{\alpha_{e_o \rightarrow \varepsilon, w}} y$	$\rho(\alpha_{e_o \rightarrow \varepsilon, w}) = \alpha$

is considered Def. 23 and Def. 24 produce the same results. Thus, in the following wherever a transformed automaton is discussed we refer to Def. 24.

Renaming  $\rho$  simply removes the extra information from the events of the transformed automaton and maps them back to the original events in the TPO. To be more specific,  $\rho$  is a map such that  $\rho(\alpha_\sigma) = \alpha$  and  $\rho(\alpha) = \alpha$ . Table I shows how the events in a transformed automaton are linked to the original events of a TPO, while the third column shows how renaming works. Specifically, in the case where events label  $\rightarrow_{yz}$  transitions, renaming does not change events names.

*Example 4:* Consider the abstracted system  $\tilde{\mathcal{G}} = \{\tilde{G}_1, \tilde{G}_2\}$ , shown in Fig. 2. The sets of secret states are  $\tilde{Q}_1^S = \{q_3\}$ ,  $\tilde{Q}_2^S = \{s_3\}$  where all the events are observable.  $T'_1$  and  $T'_2$  are the largest three-player observers of  $\tilde{G}_1$ ,  $\tilde{G}_2$ , respectively. In Example 3 the monolithic transformed automata of  $T'_1$  and  $T'_2$  were generated. The three-player observer system  $\{T'_1, T'_2\}$  is transformed to automata system  $\mathcal{G} = \{G_1^T, G_2^T\}$ , shown Fig. 2, by adding self-loops at the marked states. Event  $\beta$  is not in the alphabet of  $T'_1$  so it appears as a self-loop at all marked states in  $G_1^T$ , which correspond to  $Y$  states in  $T'_1$ . Similarly,  $\gamma$  is added as a self-loop at marked states in  $G_2^T$  since  $\gamma$  is not in the alphabet of  $T'_2$ .

In the following, Theorem 8 proves that if the synchronization of transformed individual three-player observers contains a transition, then the largest monolithic three-player observer w.r.t. the synchronized system also contains an equivalent transition. However, the inverse is not necessarily true as there are some behaviors in the monolithic three-player observer that are omitted in the modular structure. Before Theorem 8, Lemma 6 [38] and Lemma 7 establish that the modular observer and desired observer are isomorphic to their monolithic counterparts.

*Lemma 6:* [38] Let  $G_1 = \langle \Sigma_1, Q_1, \rightarrow_1, Q_1^0 \rangle$  and  $G_2 = \langle \Sigma_2, Q_2, \rightarrow_2, Q_2^0 \rangle$  be two nondeterministic automata. Then  $\det(G_1 \parallel G_2)$  is isomorphic to  $\det(G_1) \parallel \det(G_2)$ .

*Lemma 7:* Let  $G_1 = \langle Q_1, \Sigma_1, \rightarrow_1, Q_1^0 \rangle$  and  $G_2 = \langle Q_2, \Sigma_2, \rightarrow_2, Q_2^0 \rangle$  be two nondeterministic automata with sets of secret states  $Q_1^S$  and  $Q_2^S$ , respectively. Then  $\det_d(G_1) \parallel \det_d(G_2)$  is isomorphic to  $\det_d(G_1 \parallel G_2)$ .

*Proof:* From  $\det(G_1 \parallel G_2)$  is isomorphic to  $\det(G_1) \parallel \det(G_2)$  it follows that  $\det(G_1 \parallel G_2) \xrightarrow{s} X$  if and only if  $\det(G_1) \parallel \det(G_2) \xrightarrow{s} (X_1, X_2)$  and  $(x_1, x_2) \in X$  if and only if  $(x_1, x_2) \in X_1 \times X_2$ . Now we need to show that  $X \notin X_{obsd}$  if and only if  $(X_1, X_2) \notin X_{1,obsd} \times X_{2,obsd}$ .

First assume  $X \notin X_{obsd}$ , which means  $X \subseteq Q^S$ . This further

means that for all  $(x_1, x_2) \in X$ , either  $x_1 \in Q_1^S$  or  $x_2 \in Q_2^S$ , which implies either  $X_1 \notin X_{1,obsd}$  or  $X_2 \notin X_{2,obsd}$ . Thus,  $(X_1, X_2) \notin X_{1,obsd} \times X_{2,obsd}$ .

Now assume  $(X_1, X_2) \notin X_{1,obsd} \times X_{2,obsd}$ . This means either  $X_1 \notin X_{1,obsd}$  or  $X_2 \notin X_{2,obsd}$ , which implies either  $X_1 \subset Q_1^S$  or  $X_2 \subset Q_2^S$ . Hence for all  $(x_1, x_2) \in (X_1, X_2) = X$  either  $x_1 \in Q_1^S$  or  $x_2 \in Q_2^S$ , which implies  $X \subseteq Q^S$ . Thus,  $X \notin X_{obsd}$ . ■

*Theorem 8:* Let  $G_1 = \langle Q_1, \Sigma_1, \rightarrow_1, Q_1^0 \rangle$  and  $G_2 = \langle Q_2, \Sigma_2, \rightarrow_2, Q_2^0 \rangle$  be two nondeterministic automata with sets of secret states  $Q_1^S$  and  $Q_2^S$ , respectively. Let  $T_1 = \langle Q_Y^1, Q_Z^1, Q_W^1, \Sigma_1, \Sigma_1^E, \Theta_1, \rightarrow_{yz}^1, \rightarrow_{zz}^1, \rightarrow_{zw}^1, \rightarrow_{wy}^1, y_0^1 \rangle$  and  $T_2 = \langle Q_Y^2, Q_Z^2, Q_W^2, \Sigma_2, \Sigma_2^E, \Theta_2, \rightarrow_{yz}^2, \rightarrow_{zz}^2, \rightarrow_{zw}^2, \rightarrow_{wy}^2, y_0^2 \rangle$  be the largest three-player observers w.r.t.  $G_1$  and  $G_2$ , respectively. Let  $G_1^T = \langle \Sigma_{G_1^T}, Q_1, \rightarrow^1, Q_1^0, Q_1^m \rangle$  and  $G_2^T = \langle \Sigma_{G_2^T}, Q_2, \rightarrow^2, Q_2^0, Q_2^m \rangle$  be the transformed automata of  $T_1$  and  $T_2$ , respectively. Let  $T$  be the largest monolithic three-player observer w.r.t.  $G_1 \parallel G_2$ . Then let  $\rho : (\Sigma_{G_1^T} \cup \Sigma_{G_2^T}) \rightarrow (\Sigma_1 \cup \Sigma_1^E \cup \Theta_1) \cup (\Sigma_2 \cup \Sigma_2^E \cup \Theta_2)$  be a renaming. We have  $[G_1^T \parallel G_2^T \xrightarrow{s} (q_1, q_2)] \Rightarrow [T \xrightarrow{\rho(s)} q]$ .

*Proof:* In Supplementary Materials. ■

Theorem 8 shows that synchronization of individual transformed three-player observers is a subsystem of the largest monolithic three-player observer. Specifically, if there is a string  $s$  in  $G_1^T \parallel G_2^T$ , then there always exists a corresponding path  $\rho(s)$  in  $T$ . The synchronized automaton form TPOs may not always be equal to the monolithic TPO since some  $\rightarrow_{zz}$  transitions may not appear in the synchronized system. This happens when a state in the synchronization of TPOs is a combination of an original  $Z$  and an original  $Y$  state from individual TPOs, while the observable event component of the original  $Z$  state is a local event. However, as there is no difference between local and shared events in the monolithic approach of obtaining TPOs, the largest monolithic TPO contains all possible transitions of edit decisions. The proof of Theorem 8 also illustrates that for every state in the largest monolithic TPO, there exists a corresponding state in the synchronized individual TPOs in automaton form.

*Remark 2:* Although the statement of Theorem 8 concentrates on the case of two individual systems, the result can be generalized to more than two individual systems. ■

*Remark 3:* Notice that Theorem 8 illustrates that some transitions are “missing” in the synchronized automaton compared with the largest monolithic three-player observer  $T$ . It further implies that more transitions will be missing if we synchronize more individual TPOs (in automaton form). Actually, we may locate those missing transitions and add them back to the synchronized automaton  $\parallel_{i=1}^n G_i^T$ .

Specifically, consider states  $(q_1, q_2, \dots, q_n)$  and  $(q'_1, q'_2, \dots, q'_n)$  in  $\parallel_{i=1}^n G_i^T$  such that  $q_i, q'_i \in Q_Y^i \cup Q_Z^i$  for all  $i$ , i.e., every component in those states is either a  $Y$  state or a  $Z$  state from an individual transformed automaton. Then we add transition  $(q_1, q_2, \dots, q_n) \xrightarrow{\sigma} (q'_1, q'_2, \dots, q'_n)$  if there exists a set of indexes  $\mathbb{I} \in 2^{\{1, 2, \dots, n\}}$ , such that for all  $i \in \mathbb{I}$ ,  $q_i = (x_{i,d}, x_{i,f})$ ,  $q'_i = (x'_{i,d}, x'_{i,f}) \in Q_Y^i$  and  $x_{i,d} \xrightarrow{\sigma} x'_{i,d}$  in  $\det_d(G_i)$ ; while for all  $i \notin \mathbb{I}$ ,  $q_i = q'_i$ . Intuitively, the added transition implies that event  $\sigma$  may be inserted in the largest monolithic TPO w.r.t.  $\parallel_{i=1}^n G_i$ . However, due to the fact that

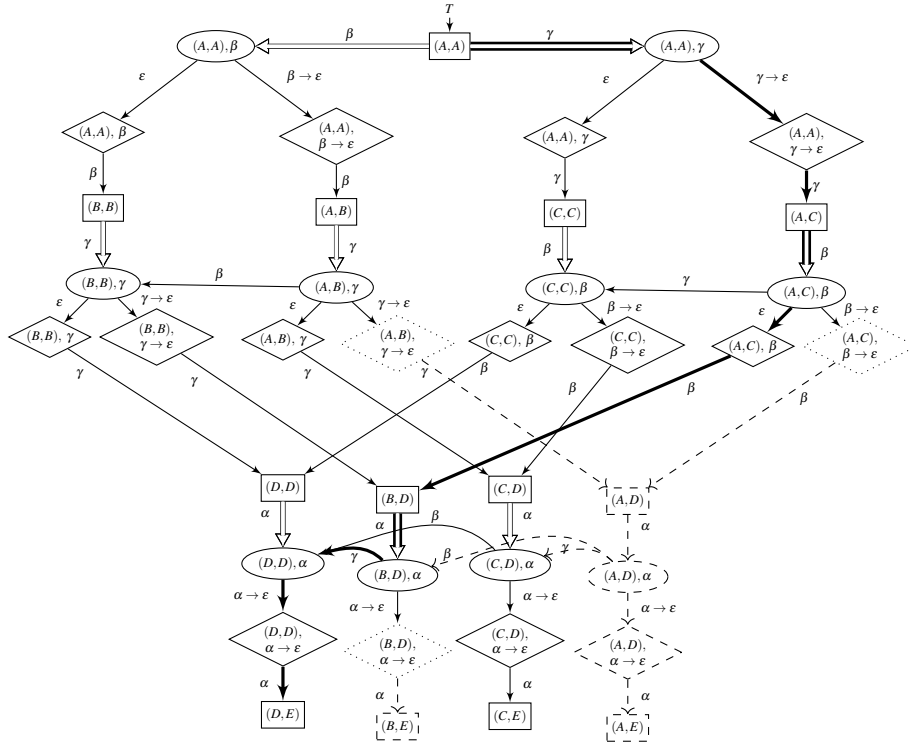


Fig. 3. The monolithic largest three-player observer w.r.t.  $G_1 \parallel G_2$  (also same as that w.r.t.  $\tilde{G}_1 \parallel \tilde{G}_2$  in this particular case) in Example 5.

there are no transitions defined from a  $Y$  state to another  $Y$  state in TPOs, those transitions are missing in  $\|_{i=1}^n G_i^T$ , which implies the synchronized system  $\|_{i=1}^n G_i^T$  may only contain a subset of edit decisions in the largest monolithic TPO.

However, the above mentioned operation may not be preferred in practice since it involves explicitly synchronizing individual TPOs in their automaton form. This is usually not feasible in modular approaches and should be avoided in our Algorithm CA-AES as well. ■

Finally, the results of this section are formally recapped in Theorem 9, which illustrates that the synchronization of transformed (automaton form) three-player observers w.r.t. individual abstracted systems contain a subset of the transitions of the largest monolithic three-player observer. The proof follows directly from Theorem 8 and Theorem 5.

**Theorem 9:** Let  $G_1$  and  $G_2$  be two nondeterministic automaton with sets of secret states  $Q_i^S \subseteq Q_i$  and sets of non-secret states  $Q_i^{NS} = Q \setminus Q_i^S$  for  $i = 1, 2$ . Let  $det(G_i)$  and  $det_d(G_i)$  be the observer and the desired observer of  $G_i$ , respectively. Let  $\sim_o$  be an opaque observation equivalence on  $G_i$  such that  $\tilde{G}_i \sim_o G_i$  for  $i = 1, 2$ . Let  $H_{i,ob} \approx_o det(\tilde{G}_i)$  and  $H_{i,b} \approx det(\tilde{G}_i)$  for  $i \in \{1, 2\}$ , where  $\approx_o$  and  $\approx$  are opaque bisimulation and bisimulation, respectively. Let  $T$  be the largest three-player observer w.r.t.  $G_1 \parallel G_2$  and let  $T_i' = \langle Q_Y^i, Q_Z^i, Q_W^i, \Sigma_i, \Sigma_i^e, \Theta_i, \rightarrow_{yz}^i, \rightarrow_{zz}^i, \rightarrow_{zw}^i, \rightarrow_{wy}^i, y_0^i \rangle$  be the largest three-player observer w.r.t.  $det_d(H_{i,ob})$  and  $H_{i,b}$  for  $i \in \{1, 2\}$ . Let  $G_i^T = \langle \Sigma_{G_i^T}, Q_i, \rightarrow_i^T, Q_i^0, Q_i^m \rangle$  for  $i = 1, 2$  be the transformed automata of  $T_i'$  and let  $\rho : (\Sigma_{G_1^T} \cup \Sigma_{G_2^T}) \rightarrow (\Sigma_1 \cup \Sigma_1^e \cup \Theta_1) \cup (\Sigma_2 \cup \Sigma_2^e \cup \Theta_2)$  be a renaming. We have  $[G_1^T \parallel G_2^T \xrightarrow{s} (q_1, q_2)] \xrightarrow{\tau} [T \xrightarrow{\rho(s)} q]$ .

**Example 5:** Consider the system  $\mathcal{G} = \{G_1, G_2\}$  shown in Fig. 2. In the first step of the compositional approach the system is abstracted by applying opaque observation equivalence, see Example 2. The abstracted system  $\tilde{\mathcal{G}} = \{\tilde{G}_1, \tilde{G}_2\}$  is shown in Fig. 2. Next, the three player observer of individual components are built. As explained in Example 4 the three-player observers of  $\tilde{G}_1$  and  $\tilde{G}_2$  are  $T_1'$  and  $T_2'$ , respectively, shown in Fig. 2. Moreover, Fig. 2 also shows  $G_1^T$  and  $G_2^T$ , the transformed automata of  $T_1'$  and  $T_2'$ , respectively. The largest monolithic three-player observer w.r.t.  $G_1 \parallel G_2$  is denoted by  $T$  and is shown in Fig. 3. In this particular example, it can be verified that the original and abstracted three-player observers are identical (this need not be true in general); therefore,  $T$  in Fig. 3 also represents the largest three-player observer w.r.t.  $\tilde{G}_1 \parallel \tilde{G}_2$ . In  $T$ , we have states:  $A = \{(q_0, s_0)\}$ ,  $B = \{(q_0, s_1), (q_0, s_2)\}$ ,  $C = \{(q_1, s_0), (q_2, s_0)\}$ ,  $D = \{(q_1, s_1), (q_2, s_1), (q_2, s_2), (q_1, s_2)\}$  and  $E = \{(q_3, s_3)\}$ .

After synchronizing  $G_1^T$  with  $G_2^T$ , we find that there are some transitions in Fig. 3, which do not correspond to any transition in  $G_1^T \parallel G_2^T$ . For example, no transition in  $G_1^T \parallel G_2^T$  corresponds to the  $zz$  transition of  $\beta$  from state  $(A, B, \gamma)$  to  $(B, B, \gamma)$  in Fig. 3.

## V. FROM ALL EDIT STRUCTURE CALCULATION TO SUPERVISOR SYNTHESIS

So far we have shown that in our compositional and abstraction-based approach, individual components can be abstracted and each largest three-player observer w.r.t. an abstracted component can be calculated individually. Then we transfer those three-player observers (TPOs) to their automaton forms. After that, we have also shown in Theorem 9 that

the synchronization of the transformed three-player observers results in a subsystem of the largest monolithic three-player observer up to the renaming of events.

Recall that the All Edit Structure (AES) is obtained after pruning deadlocking states from the largest TPO. Here the modular structure of the transformed TPO is kept and the calculation of a “Modular Edit Structure” can be done by mapping this problem to a modular nonblocking supervisory control problem under full observation.

As was discussed at the end of Section II-C, we pursue an approach to convert the pruning process (from the largest TPO to the AES) to a supervisory control problem. In this setting, the plant is a collection of automata transformed from individual largest TPOs obtained at the end of step (iv) of Algorithm CA-AES. The specification is the automaton form of the edit constraint. The constraint of having up to  $n + 1$  consecutive erasures can be modeled by a specification automaton with  $n$  states where transitions are labeled by the decision events and all states are marked except the last state, which is a blocking state. After  $n$  consecutive event erasures, the next transition of event erasure  $\alpha \rightarrow \varepsilon_\alpha$  leads the specification forward to a blocking state. If the next event is a non-erasure event, it leads the specification back to the initial state, thus resetting the sequence of erasures. Since we have a *modular* representation of the plant, we are able to leverage computationally efficient compositional techniques for modular nonblocking supervisory control problems.

**Definition 25:** Let  $\mathcal{T} = \{T_1, \dots, T_n\}$  be a three-player observer system where  $T_i = \langle Q_Y^i, Q_Z^i, Q_W^i, \Sigma^i, \Sigma_\varepsilon^i, \Theta^i, \rightarrow_{yz}^i, \rightarrow_{zz}^i, \rightarrow_{zw}^i, \rightarrow_{wy}^i, y_0^i \rangle$  and let  $\Phi$  be the edit constraint on  $\mathcal{T}$  such that there are not  $n + 1$  consecutive event erasures. Then  $K = \langle \Sigma_K, Q_K, \rightarrow_K, Q_K^0, Q_K^m \rangle$  is the automaton form of  $\Phi$  where,

- $Q_K = \{x_1, \dots, x_n\}$
- $\rightarrow_K = \bigcup_{1 \leq i \leq n-1} \{(x_i, \alpha \rightarrow \varepsilon_\alpha, x_{i+1}) \mid p \xrightarrow{\alpha \rightarrow \varepsilon} z_{w2} q \text{ and } E(p) = \alpha\} \cup \bigcup_{1 \leq i \leq n-2} \{(x_{i+1}, \varepsilon_\alpha, x_1) \mid p \xrightarrow{\varepsilon} z_{w1} q \text{ and } E(p) = \alpha\} \cup \bigcup_{1 \leq i \leq n-2} \{(x_{i+1}, \alpha_\sigma, x_1) \mid p \xrightarrow{\alpha} z_{zz} q \text{ and } E(p) = \sigma\} \cup \{(x_1, \varepsilon_\alpha, x_1) \mid p \xrightarrow{\varepsilon} z_{w1} q \text{ and } E(p) = \alpha\} \cup \{(x_1, \alpha_\sigma, x_1) \mid p \xrightarrow{\alpha} z_{zz} q \text{ and } E(p) = \sigma\},$
- $Q_K^0 = x_1$
- $Q_K^m = \{x_1, \dots, x_{n-1}\}$

**Example 6:** Consider the transformed system  $\mathcal{G}^T = \{G_1^T, G_2^T\}$  shown in Fig. 2. Assume the constraint  $\Phi$  only allows one erasure. The specification automaton of this constraint is shown in Fig. 4 as  $K$ . Automaton  $K$  has three states. As there is no constraint on event insertion, the events related with event insertion just form self-loops at the initial state of  $K$ . On the other hand, by executing  $\alpha \rightarrow \varepsilon_\alpha$ ,  $\beta \rightarrow \varepsilon_\beta$  or  $\gamma \rightarrow \varepsilon_\gamma$  the specification transits from  $x_1$  to  $x_2$ . Next, at  $x_2$  if the edit decision is to certain events, then the system goes back to the initial state  $x_1$ , thus allowing more event erasures since there are no consecutive erasures. However, if another event erasure occurs from  $x_2$ , the system goes to the blocking state  $x_3$ .

The following theorem establishes that the three-player observer of the system under constraint  $\Phi$  and the transformed system synchronized with the specification  $K$  have the same runs up to a renaming of the events.

**Theorem 10:** Let  $G = \langle \Sigma, Q, \rightarrow, Q^0 \rangle$  be a nondeterministic

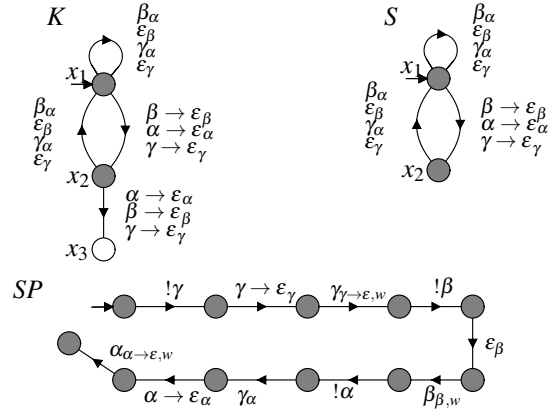


Fig. 4. Automaton  $K$  is the automaton form of the constraint  $\Phi$  in Example 6,  $S$  and  $SP$  are the supervisor and the selected path in Example 7

automaton with the set of secret states  $Q^S \subseteq Q$ . Let  $T' = \langle Q_Y', Q_Z', Q_W', \Sigma, \Sigma_\varepsilon, \Theta, \rightarrow_{yz}', \rightarrow_{zz}', \rightarrow_{zw}', \rightarrow_{wy}', y_0' \rangle$  be the largest three-player observer of  $G$  under the edit constraint  $\Phi$  which prohibits  $n + 1$  consecutive event erasures. Let  $T = \langle Q_Y, Q_Z, Q_W, \Sigma, \Sigma_\varepsilon, \Theta, \rightarrow_{yz}, \rightarrow_{zz}, \rightarrow_{zw}, \rightarrow_{wy}, y_0 \rangle$  be the largest three-player observer w.r.t.  $G$  when no constraint is considered and let  $G^T = \langle \Sigma_{G^T}, Q, \rightarrow, Q^0 \rangle$  be the automaton transformation of  $T$ . Let  $K$  be the specification automaton of  $\Phi$ . Then  $G^T \parallel K \xrightarrow{s}$  if and only if  $T' \xrightarrow{\rho(s)}$ .

*Proof:* Clearly  $T' \subseteq T$  and  $\Sigma_K \subseteq \Sigma_{G^T}$ . First let  $G^T \parallel K \xrightarrow{s} (q_G, q_K)$ , let  $P_K : \Sigma_{G^T} \rightarrow \Sigma_K$  and let  $P_E : \Sigma_{G^T} \rightarrow \bigcup_{\sigma \in \Sigma} \{\sigma \rightarrow \varepsilon_\sigma \in \Sigma_K\}$  be a map that removes all the events except event erasures from  $\Sigma_K$ . From  $G^T \parallel K \xrightarrow{s} (q_G, q_K) \xrightarrow{\sigma} (p_G, p_K)$  and  $\Sigma_K \subseteq \Sigma_{G^T}$  and Def. 2, it holds that  $G^T \xrightarrow{s} q_G \xrightarrow{\sigma} p_G$  and  $K \xrightarrow{P_K(s)} q_K \xrightarrow{P_K(\sigma)} p_K$ . Now consider three cases:

- 1)  $q_K = x_i$ ,  $p_K = x_{i+1}$  and  $p_K, q_K \in Q_m^K$ . Then  $|P_E(s)| < n$ . This implies  $\sigma$  is an erasure event but there are not  $n$  consecutive erasures in  $s$ , so  $T' \xrightarrow{\rho(s)} q_T \xrightarrow{\rho(\sigma)} p_T$ .
- 2)  $q_K = x_i$ ,  $p_K = x_{i+1}$ ,  $q_K \in Q_m^K$  but  $p_K \notin Q_m^K$ . Then  $|P_E(s)| = n$  and  $K \xrightarrow{P_K(s)} q_K \xrightarrow{\sigma} p_K$ , which implies  $G^T \parallel K \xrightarrow{s} (p_G, p_K)$  and  $(p_G, p_K)$  is a blocking state. This further indicates there are  $n$  consecutive erasures in  $s$ , so  $T' \xrightarrow{\rho(s)} q_T \not\xrightarrow{\rho(\sigma)}$ .
- 3)  $q_K = x_i$ ,  $p_K = x_1$  and  $p_K, q_K \in Q_m^K$ . This implies  $\sigma$  is a non-erasure event, so  $T' \xrightarrow{\rho(s)} q_T \xrightarrow{\rho(\sigma)} p_T$ .

Now assume  $T' \xrightarrow{\rho(s)} q_T$ . This means  $T \xrightarrow{\rho(s)} q_T$ , which implies  $G^T \xrightarrow{s} q_T$ . Consider two cases. 1.  $|P_E(s)| < n$ . Then  $K \xrightarrow{P_K(s)} q_K$ , which implies  $G^T \parallel K \xrightarrow{s} (q_G, q_K)$ . 2.  $|P_E(s)| = n$ . This means  $T' \xrightarrow{\rho(s)} q_T \not\xrightarrow{\rho(\sigma)}$  and  $K \xrightarrow{P_K(s)} q_K \not\xrightarrow{\rho(\sigma)}$ , which implies  $G^T \parallel K \xrightarrow{s} (q_G, q_K) \not\xrightarrow{\rho(\sigma)}$ . ■

The following theorem shows that equivalent states are removed in solving the supervisory control problem to obtain an automaton satisfying the edit constraint and in the pruning process to obtain the AES from the largest TPO.

**Theorem 11:** Let  $G = \langle \Sigma, Q, \rightarrow, Q^0 \rangle$  be a nondeterministic automaton with the set of secret states  $Q^S \subseteq Q$  and let  $\Phi$  be the edit constraint which prohibits  $n + 1$  consecutive event erasures. Let  $T = \langle Q_Y, Q_Z, Q_W, \Sigma, \Sigma_\varepsilon, \Theta, \rightarrow_{yz}, \rightarrow_{zz}, \rightarrow_{zw}, \rightarrow_{wy}, y_0 \rangle$

be the largest three-player observer w.r.t.  $G$  without considering the edit constraint and let  $G^T = \langle \Sigma_{G^T}, Q, \rightarrow, Q^0 \rangle$  be the transformed automaton of  $T$ . Let  $\rho : \Sigma_{G^T} \rightarrow (\Sigma_1 \cup \Sigma_1^\varepsilon \cup \Theta_1)$  be a renaming and let  $AES$  be the All Edit Structure obtained from  $T$ . Let  $S$  be the supremal controllable and nonblocking subautomaton of  $G^T$  after considering the specification introduced by  $\Phi$ . Then  $S \xrightarrow{\rho^{-1}(s)} q$  if and only if  $AES \xrightarrow{\rho(s)} q$ .

*Proof:* In Supplementary Materials. ■

Theorem 11 proves that when it comes to imposing the edit constraint, the pruning process from the largest TPO to the AES removes equivalent states with the synthesis procedure of a supremal supervisor. Hence no information is lost when we apply the supervisory control approach to enforce the edit constraints and obtain edit functions. This result is essential to show that the transformation of the TPO to an equivalent automaton and the constraint  $\Phi$  to specification  $K$  is correctly done in Def. 23 and Def. 25, respectively. The next step is to consider the modular representation of the system. In that case, we will use the transformation in Def. 24, which results in a set of automata transformations of the individual three-player observers, with necessary self-loops to capture the synchronization among the components. Finally, we combine the results about abstraction and decomposition, which results in Theorem 12.

**Theorem 12:** Let  $\mathcal{G} = \{G_1, \dots, G_n\}$  be a modular nondeterministic system with sets of secret states  $Q_i^S$ . Let  $AES$  be the All Edit Structure of  $\mathcal{G}$  under constraint  $\Phi$ . Let  $det(G_i)$  and  $det_d(G_i)$  be the observer and the desired observer of  $G_i$ , respectively. Let  $\sim_o$  be an opaque observation equivalence on  $G_i$  such that  $\tilde{G}_i \sim_o G_i$  for  $i = 1, \dots, n$ . Let  $H_{i,ob} \approx_o det(\tilde{G}_i)$  and  $H_{i,b} \approx det(\tilde{G}_i)$  for  $i = 1, \dots, n$ , where  $\approx_o$  and  $\approx$  are opaque bisimulation and bisimulation, respectively. Let  $T'_i$  be the largest three-player observer of  $det_d(H_{i,ob})$  and  $H_{i,b}$  for  $i = 1, \dots, n$  with the event set  $\Sigma_{i,T}$ . Let  $G_i^T$  be the transformed automaton of  $T'_i$  and  $K$  be the automaton specification. Let  $P_e : \Omega \rightarrow \mathcal{L}(\mathcal{G})$  be an edit projection and  $l(\omega)$  be a string generated by run (Def. 15 and Def. 14). Let  $\mathcal{S}$  be the least restrictive controllable and nonblocking supervisor calculated from  $\{G_1^T, \dots, G_n^T, K\}$  and let  $\rho : (\Sigma_{G_1^T} \cup \dots \cup \Sigma_{G_n^T}) \rightarrow (\Sigma_{1,T} \cup \dots \cup \Sigma_{n,T})$  be the renaming map. Then  $[\forall s \in \mathcal{L}(\mathcal{G}), \exists t \in \mathcal{L}(\mathcal{S}): P_e(\rho(t)) = s] \Rightarrow [l(\rho(t)) = f_e(s) \text{ where } f_e \in AES]$ .

*Proof:* The proof follows directly from Theorems 11 and 10, in combination with Theorem 9. ■

Theorem 12 essentially shows the proof for all the steps shown in Fig. 1. The theorem shows that Algorithm CA-AES correctly synthesizes edit functions for opacity enforcement in a modular form, therefore, the algorithm is sound. It also reveals that the problem of calculating the modular representation of the All Edit Structure can be transformed to synthesizing modular supervisors. The advantage of such a transformation is that we may leverage various existing approaches for calculating a modular supremal nonblocking supervisor in the literature; see, e.g., [11], [28], [29], [37]. Therefore, we can obtain a modular representation of the All Edit Structure, which is noticeably efficient to compute. Then we may synchronize individual components in the Modular Edit Structure, which results in a subsystem of the monolithic

AES. However, as was pointed out in Section IV, some edit decisions are omitted after the synchronization. In practice, it is usually desired to retain the modular structure and extract an edit function from it, much in the same way as a set of modular supervisors control a plant. The extraction process is explained next.

Each step of extracting a valid edit decision is described in Fig. 5. Here the edit function is an interface between the system's output and the outside environment. Assume that the system outputs event  $\gamma$ , then the edit function makes an edit decision for that event and the edited string will be output to the external observers.

Specifically, this process contains the following steps. (1) when  $\gamma$  is received by the edit function, all the components of the Modular Edit Structure are in states that correspond to  $Y$  states of the All Edit Structure. (2) At these states, event  $\gamma$  is executed and states of all the components in the Modular Edit Structure are updated simultaneously. After the execution of  $\gamma$ , each component of the Modular Edit Structure is at a state that corresponds to a  $Z$  state of the All Edit Structure. (3) Then assume there are multiple transitions defined out of such a current state, we need to select one common transition which corresponds to a specific edit decision and can be viewed as making a control decision from the current state. Note that as the selected transition needs to be accepted by all the components, thus it may happen spontaneously in all the components of the system. The solution of the modular supervisory control problem guarantees the existence of such a common transition out of the current  $Z$  states.

Algorithm CA-AES returns the edited string  $\rho(\sigma_0 \dots \sigma_k)$  for event  $\gamma$  when every component of the Modular Edit Structure reaches a new state corresponding to a  $Y$  state of the AES. At that point, the Modular Edit Structure is ready to process the next event output by the system, and the above steps repeat. Meanwhile, the algorithm keeps track of the states of the Modular Edit Structure as its components evolve. Based on Theorem 11, the edited string  $\rho(\sigma_0 \dots \sigma_k)$  is accepted by the monolithic AES. This finally confirms that the extracted edit decision from the Modular Edit Structure corresponds a valid edit decision in the monolithic AES. The above process is illustrated in the following example.

**Example 7:** Consider the nondeterministic system  $\mathcal{G} = \{G_1, G_2\}$  shown in Fig. 2. As it was shown in Example 2, the system can be abstracted using opaque observation equivalence. After abstraction, the system becomes deterministic, which means there is no need to calculate the observers of  $G_1$  and  $G_2$ . The largest three-player observers of  $\tilde{G}_1$  and  $\tilde{G}_2$  are  $T'_1$  and  $T'_2$ , respectively, shown in Fig. 2. Next, the three-player observers are transformed to automata  $G_1^T$  and  $G_2^T$  shown in Fig. 2, as explained in Example 4.

Assume the user adds an edit constraint such that only one consecutive erasure is allowed as in Example 6. The specification automaton of this constraint is  $K$ , shown in Fig. 4. Due to this constraint, the  $Y$  states  $(A, D)$  and  $(B, E)$  are considered undesired states in  $T$ , shown in Fig. 3 and they should not be reached when we synthesize edit functions. Since  $(A, D)$  is not allowed, its successor states  $(A, D, \alpha)$ ,  $(A, D, \alpha \rightarrow \varepsilon)$  and  $(A, E)$  become unreachable from the initial

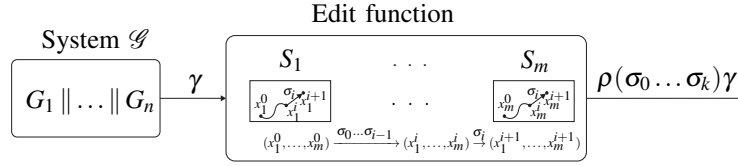


Fig. 5. Process of selecting an edit decision from the modular All Edit Structure.

state  $(A, A)$ . Those three states together with  $(A, D)$  and  $(B, E)$  are drawn in dashed lines in Fig. 3 and are to be removed in the next step. Furthermore, states  $(B, D, \alpha \rightarrow \varepsilon)$ ,  $(A, C, \beta \rightarrow \varepsilon)$  and  $(A, B, \gamma \rightarrow \varepsilon)$  become deadlocking after  $(A, D)$  and  $(B, E)$  are removed. They are drawn in dotted lines in Fig. 3 and are also to be removed.

Following the compositional approach with supervisor reduction presented in [29], [40], we calculate a least restrictive and nonblocking supervisor for the transformed automaton, which is shown in Fig. 4 as automaton  $S$ . All the paths accepted by this supervisor represent valid edit decisions. Consider the accepted path  $SP$  shown in Fig. 4, it corresponds to an edit function's decisions for string  $\gamma\beta\alpha$  such that  $f_e(\gamma\beta\alpha) = \gamma(\gamma \rightarrow \varepsilon)\gamma\beta\varepsilon\beta\alpha\gamma(\alpha \rightarrow \varepsilon)\alpha$ , which is shown by thick lines in  $T$ , Fig. 3. As is seen,  $\rho(SP) = f_e(\gamma\beta\alpha)$ . Specifically, when event  $\gamma$  is output by the system,  $SP$  returns  $\gamma(\gamma \rightarrow \varepsilon_\gamma)(\gamma_{\gamma \rightarrow \varepsilon, w})$ , which means erasing the  $\gamma$ . Next, event  $\beta$  is output by the system and it is unchanged according to  $SP$ . Finally,  $\alpha$  is output by the system and  $SP$  returns  $\alpha(\gamma_\alpha)(\alpha \rightarrow \varepsilon_\alpha)(\alpha_{\alpha \rightarrow \varepsilon, w})$ , which means erasing  $\alpha$  and inserting  $\gamma$ . Similarly, we may consider other paths accepted by the supervisor in Fig. 4 to track edit decisions on other strings, then we have a complete picture of how an edit function works.

*Remark 4:* From the result in Theorem 8, our modular algorithm CA-AES results in fewer edit decisions compared with the monolithic approach in [20], due to the synchronization process in Section IV. This indicates that our method may not be complete in the sense that even if Algorithm CA-AES does not return any modular form edit functions, the monolithic approach may still return valid edit functions. This may be viewed as the tradeoff of reducing computational complexity by the modular method. ■

## VI. CONCLUSION

This paper investigated a compositional and abstraction-based approach to synthesize edit functions for opacity enforcement in a modular setting, given a set of individual systems. The edit functions modify the system's output by inserting and erasing events, under the constraint of limited number of event erasures. The Three-Player Observer (TPO) and All Edit Structure (AES) proposed in our prior work were employed here; these discrete structures embed edit functions and reflect the constraints. The monolithic approach first synchronizes all individual systems, then calculates the monolithic AES to obtain edit functions. In contrast, the compositional approach first exploits the modular structure and builds individual TPOs. Then, it incorporates the edit constraint and calculates the Modular Edit Structure in a nonblocking modular supervisory control manner to obtain edit functions. In addition, we also applied abstraction methods

to reduce the state space of the system before opacity enforcement. We showed that the abstraction processes preserve opacity. Combining system composition and abstraction, we proposed an efficient approach to enforce opacity for complex systems containing multiple components.

## REFERENCES

- [1] B. Bérard, K. Chatterjee, and N. Sznajder. Probabilistic opacity for Markov decision processes. *Information Processing Letters*, 115(1):52–59, 2015.
- [2] J. W. Bryans, M. Koutny, L. Mazaré, and P. Y. A. Ryan. Opacity generalised to transition systems. *International Journal of Information Security*, 7(6):421–435, 2008.
- [3] J. W. Bryans, M. Koutny, and P.Y.A. Ryan. Modelling opacity using Petri nets. *Elec. Notes in Theor. Computer Science*, 121:101–115, 2005.
- [4] C. G. Cassandras and S. Lafortune. *Introduction to discrete event systems – 2nd Edition*. Springer, 2008.
- [5] F. Cassez, J. Dubreil, and H. Marchand. Synthesis of opaque systems with static and dynamic masks. *Formal Methods in System Design*, 40(1):88–115, 2012.
- [6] S. Chédor, C. Morvan, S. Pinchinat, and H. Marchand. Diagnosis and opacity problems for infinite state systems modeled by recursive tile systems. *Discrete Event Dynamic Systems: Theory and Applications*, 25(1-2):271–294, 2015.
- [7] J. Chen, M. Ibrahim, and R. Kumar. Quantification of secrecy in partially observed stochastic discrete event systems. *IEEE Transactions on Automation Science and Engineering*, 14(1):185–195, 2017.
- [8] P. Darondeau, H. Marchand, and L. Ricker. Enforcing opacity of regular predicates on modal transition systems. *Discrete Event Dynamic Systems: Theory and Applications*, 25(1-2):251–270, 2015.
- [9] J. Dubreil, P. Darondeau, and H. Marchand. Supervisory control for opacity. *IEEE Trans. on Automatic Control*, 55(5):1089–1100, 2010.
- [10] Y. Falcone and H. Marchand. Enforcement and validation (at runtime) of various notions of opacity. *Discrete Event Dynamic Systems: Theory and Applications*, 25(4):531–570, 2015.
- [11] L. Feng and W. M. Wonham. Supervisory control architecture for discrete-event systems. *IEEE Transactions on Automatic Control*, 53(6):1449–1461, 2008.
- [12] J.-C. Fernandez. An implementation of an efficient algorithm for bisimulation equivalence. *Science of Computer Programming*, 13(2-3):219–236, 1990.
- [13] H. Flordal, R. Malik, M. Fabian, and K. Åkesson. Compositional synthesis of maximally permissive supervisors using supervision equivalence. *Discrete Event Dynamic Systems: Theory and Applications*, 17(4):475–504, 2007.
- [14] L. Hélouët, H. Marchand, and L. Ricker. Opacity with powerful attackers. In *Proceedings of the 14th IFAC International Workshop on Discrete Event Systems*, pages 475–482, 2018.
- [15] C. A. R. Hoare. Communicating sequential processes. *Communications of the ACM*, 21(8):666–677, 1978.
- [16] R. Jacob, J.-J. Lesage, and J.-M. Faure. Overview of discrete event systems opacity: Models, validation, and quantification. *Annual Reviews in Control*, 2016.
- [17] Y. Ji and S. Lafortune. Enforcing opacity by publicly known edit functions. In *Proceedings of the 56th IEEE Conference on Decision and Control*, pages 4866–4871, 2017.
- [18] Y. Ji, Y.-C. Wu, and S. Lafortune. Enforcement of opacity by public and private insertion functions. *Automatica*, 93:369–378, 2018.
- [19] Y. Ji, Yin. X, and S. Lafortune. Enforcing opacity by insertion functions under multiple energy constraints. *Automatica*, accepted, 2019.
- [20] Y. Ji, X. Yin, and S. Lafortune. Opacity enforcement using nondeterministic publicly-known edit functions. *IEEE Transactions on Automatic Control*, DOI: 10.1109/TAC.2019.2897553, to appear, 2019.



- [21] C. Keroglou and C. N. Hadjicostis. Probabilistic system opacity in discrete event systems. *Discrete Event Dynamic Systems: Theory and Applications*, 28(2):289–314, 2018.
- [22] F. Lin. Opacity of discrete event systems and its applications. *Automatica*, 47(3):496–503, 2011.
- [23] R. Malik and R. Leduc. Compositional nonblocking verification using generalized nonblocking abstractions. *IEEE Transactions on Automatic Control*, 58(8):1891–1903, 2013.
- [24] T. Masopust and X. Yin. Complexity of detectability, opacity and A-diagnosability for modular discrete event systems. *Automatica*, 101:290–295, 2019.
- [25] R. Milner. *Communication and Concurrency*, volume 84. Prentice Hall New York, 1989.
- [26] S. Mohajerani, Y. Ji, and S. Lafortune. Efficient synthesis of edit functions for opacity enforcement using bisimulation-based abstractions. In *Proceedings of the 57th IEEE Conference on Decision and Control*, pages 3573–3578, 2018.
- [27] S. Mohajerani and S. Lafortune. Transforming opacity verification to nonblocking verification in modular systems. *IEEE Transactions on Automatic Control*, 2018, under review.
- [28] S. Mohajerani, R. Malik, and M. Fabian. A framework for compositional synthesis of modular nonblocking supervisors. *IEEE Transactions on Automatic Control*, 59(1):150–162, 2014.
- [29] S. Mohajerani, R. Malik, and M. Fabian. Compositional synthesis of supervisors in the form of state machines and state maps. *Automatica*, 76:277–281, 2017.
- [30] J. Mullins and M. Yeddes. Opacity with orwellian observers and intran-sitive non-interference. In *Proceedings of the 12th IFAC International Workshop on Discrete Event Systems*, pages 344–349, 2014.
- [31] M. Noori-Hosseini, B. Lennartson, and C. Hadjicostis. Compositional visible bisimulation abstraction applied to opacity verification. In *Proceedings of the 14th IFAC International Workshop on Discrete Event Systems*, pages 434–441, 2018.
- [32] J. JMM Rutten. Automata and coinduction (an exercise in coalgebra). In *Intl. Conference on Concurrency Theory*, pages 194–218, 1998.
- [33] A. Saboori and C. N. Hadjicostis. Verification of infinite-step opacity and complexity considerations. *IEEE Transactions on Automatic Control*, 57(5):1265–1269, 2012.
- [34] A. Saboori and C.N. Hadjicostis. Notions of security and opacity in discrete event systems. In *Proceedings of the 46th IEEE Conference on Decision and Control*, pages 5056–5061. IEEE, 2007.
- [35] A. Saboori and C.N. Hadjicostis. Opacity-enforcing supervisory strategies via state estimator constructions. *IEEE Transactions on Automatic Control*, 57(5):1155–1165, 2012.
- [36] A. Saboori and C.N. Hadjicostis. Verification of initial-state opacity in security applications of discrete event systems. *Information Sciences*, 246:115–132, 2013.
- [37] K. W. Schmidt and J. Eduardo R. Cury. Efficient abstractions for the supervisory control of modular discrete event systems. *IEEE Transactions on Automatic Control*, 57(12):3224–3229, 2012.
- [38] C. Seatzu, M. Silva, and J. H. van Schuppen. *Control of Discrete-Event Systems: Automata and Petri Net Perspectives*. SPRINGER, 2012.
- [39] R. Su, J. H. van Schuppen, and J. E. Rooda. Model abstraction of nondeterministic finite-state automata in supervisor synthesis. *IEEE Transactions on automatic control*, 55(11):2527–2541, 2010.
- [40] R. Su and W. M. Wonham. Supervisor reduction for discrete-event systems. *Discrete Event Dynamic Systems: Theory and Applications*, 14(1):31–53, 2004.
- [41] S. Takai and Y. Oka. A formula for the supremal controllable and opaque sublanguage arising in supervisory control. *SICE Journal of Control, Measurement, and System Integration*, 1(4):307–311, 2008.
- [42] Y. Tong, Z. Li, C. Seatzu, and A. Giua. Decidability of opacity verification problems in labeled petri net systems. *Automatica*, 80:48–53, 2017.
- [43] Y. Tong, Z. Li, C. Seatzu, and A. Giua. Verification of state-based opacity using Petri nets. *IEEE Transactions on Automatic Control*, 62(6):2823–2837, 2017.
- [44] W. M. Wonham and K. Cai. *Supervisory Control of Discrete-Event Systems*. SPRINGER, 2019.
- [45] B. Wu and H. Lin. Privacy verification and enforcement via belief abstraction. *IEEE control systems letters*, 2(4):815–820, 2018.
- [46] Y.-C. Wu and S. Lafortune. Synthesis of insertion functions for enforcement of opacity security properties. *Automatica*, 50(5):1336–1348, 2014.
- [47] Y.-C. Wu, V. Raman, B. C. Rawlings, S. Lafortune, and S. A. Seshia. Synthesis of obfuscation policies to ensure privacy and utility. *Journal of Automated Reasoning*, 60(1):107–131, 2018.
- [48] X. Yin and S. Lafortune. A uniform approach for synthesizing property-enforcing supervisors for partially-observed discrete-event systems. *IEEE Transactions on Automatic Control*, 61(8):2140–2154, 2016.
- [49] X. Yin and S. Lafortune. A new approach for the verification of infinite-step and K-step opacity using two-way observers. *Automatica*, 80:162–171, 2017.
- [50] X. Yin and S. Lafortune. A general approach for optimizing dynamic sensor activations for discrete event systems. *Automatica*, 105:376–383, 2019.
- [51] X. Yin and S. Li. Verification of opacity in networked supervisory control systems with insecure control channels. In *Proceedings of 57th IEEE Conference on Decision and Control*, pages 4851–4856, 2018.
- [52] X. Yin, Z. Li, W. Wang, and S. Li. Infinite-step opacity and K-step opacity of stochastic discrete-event systems. *Automatica*, 99:266–274, 2019.
- [53] B. Zhang, S. Shu, and F. Lin. Maximum information release while ensuring opacity in discrete event systems. *IEEE Transactions on Automation Science and Engineering*, 12(3):1067–1079, 2015.
- [54] K. Zhang, X. Yin, and M. Zamani. Opacity of nondeterministic transition systems: A (bi) simulation relation approach. *IEEE Transactions on Automatic Control*, DOI: 10.1109/TAC.2019.2908726, to appear, 2019.



event systems and cyber physical systems.

**Sahar Mohajerani** received her B.Sc. degree in electrical engineering from Khaje Nasir Toosi University of Technology in 2005, the M.S. degree in systems, control, and mechatronics from Chalmers University of Technology in 2009 and her Ph.D. degree in automation from Chalmers University of Technology in 2015. She worked at Volvo Cars Corporation on function verification. She is currently a post-doctoral fellow at the University of Michigan. Her research interests are in the area of formal methods for verification and synthesis of large discrete



tems.

**Yiding Ji** received the Bachelor of Engineering degree of Electrical Engineering from Tianjin University, Tianjin, China, in 2014, the Master of Science degree and PhD degree of Electrical and Computer Engineering from the University of Michigan, USA, in 2016 and 2019, respectively.

His research interests include theory of discrete event systems, security and privacy in cyber and cyber-physical systems, formal methods and game theory. He is a member of the IEEE Control Systems Society Technical Committee on Discrete Event Sys-



**Stéphane Lafortune** received the B.Eng degree from École Polytechnique de Montréal in 1980, the M.Eng degree from McGill University in 1982, and the Ph.D degree from the University of California at Berkeley in 1986, all in electrical engineering. Since September 1986, he has been with the University of Michigan, Ann Arbor, where he is a Professor of Electrical Engineering and Computer Science. Dr. Lafortune is a Fellow of the IEEE (1999) and of IFAC (2017). He received the Presidential Young Investigator Award from the National Science Foundation in 1990 and the Axelby Outstanding Paper Award from the Control Systems Society of the IEEE in 1994 (for a paper co-authored with S.-L. Chung and F. Lin) and in 2001 (for a paper co-authored with G. Barrett). Dr. Lafortune's research interests are in discrete event systems and include multiple problem domains: modeling, diagnosis, control, optimization, and applications to computer and software systems.

## Supplementary Materials

**Theorem 5:** Let  $G$  be a nondeterministic automaton with secret states  $Q^S \subseteq Q$  and non-secret states  $Q^{NS} = Q \setminus Q^S$ . Let  $\det(G)$  and  $\det_d(G)$  be the observer and the desired observer of  $G$ , respectively. Let  $\sim_o$  be an opaque observation equivalence on  $G$  such that  $\tilde{G} \sim_o G$ . Let  $H_{ob} \approx_o \det(\tilde{G})$  and  $H_b \approx \det(\tilde{G})$  where  $\approx_o$  and  $\approx$  are opaque bisimulation and bisimulation, respectively. Let  $H_{obd}$  be the desired observer of  $H_{ob}$ . Let  $T$  be the largest three-player observer w.r.t.  $\det(G)$  and  $\det_d(G)$ , also let  $T'$  be the largest three-player observer w.r.t.  $H_{obd}$  and  $H_b$ . Then  $T \xrightarrow{\omega} q$  if and only if  $T' \xrightarrow{\omega} \tilde{q}$ .

*Proof:* From Propositions 3 and 4 it holds that  $\det(G) \approx H_b$  and  $\det_d(G) \approx H_{obd}$ . Thus, we need to show that a transition is defined in  $T$  if and only if the same transition is defined in  $T'$ . It is shown by induction on  $n \geq 0$  that  $y^0 \xrightarrow{\omega} q_n$  in  $T$  if and only if  $\tilde{y}^0 \xrightarrow{\omega} \tilde{q}_n$  in  $T'$ .

*Base case:* ( $\Rightarrow$ ) First assume  $y^0 \xrightarrow{e_a} z^0$  in  $T$ , where  $y^0 = (X_{obsd}^0, X_{obs}^0)$ . Based on Def. 12 it holds that  $X_{obs}^0 \xrightarrow{e_a}$  in  $\det(G)$ ,  $I(z^0) = y^0$  and  $E(z^0) = e_o$ . From  $X_{obs}^0 \xrightarrow{e_a}$  in  $\det(G)$  and since  $\det(G) \approx H_b$  it holds that  $\tilde{X}_{obs}^0 \xrightarrow{e_a}$  in  $H_b$ . Thus,  $\tilde{y}^0 = (\tilde{X}_{obsd}^0, \tilde{X}_{obs}^0)$  and  $\tilde{X}_{obs}^0 \xrightarrow{e_a}$  in  $H_b$ ,  $I(\tilde{z}^0) = \tilde{y}^0$  and  $E(\tilde{z}^0) = e_o$ . This means  $\tilde{y}^0 \xrightarrow{e_a} \tilde{z}^0$  in  $T'$ .

( $\Leftarrow$ ) Now assume  $\tilde{y}^0 \xrightarrow{e_a} \tilde{z}^0$  in  $T'$ , where  $\tilde{y}^0 = (\tilde{X}_{obsd}^0, \tilde{X}_{obs}^0)$ . The same argument as ( $\Rightarrow$ ) holds.

*Inductive step:* Assume the claim holds for some  $n \geq 0$ , i.e.,  $n \geq 0$  that  $y^0 \xrightarrow{\omega} q_n$  in  $T$  if and only if  $\tilde{y}^0 \xrightarrow{\omega} \tilde{q}_n$  in  $T'$ .

( $\Rightarrow$ ) It must be shown that if  $q_n \xrightarrow{\sigma} p_n$  in  $T$  then  $\tilde{q}_n \xrightarrow{\sigma} \tilde{p}_n$  in  $T'$ . There are six possibilities:

- $q_n = y$  is a  $Y$  state and  $p_n = z$  is a  $Z$  state, Def. 12 (i). Let  $y = (x_d, x_f)$  and  $\sigma \in \Sigma$ . Based on inductive assumption there exists  $\tilde{y} = (\tilde{x}_d, \tilde{x}_f)$  and  $\tilde{y}^0 \xrightarrow{\omega} \tilde{y}$  in  $T'$ . From  $y \xrightarrow{\sigma} z$  in  $T$  and based on Def. 12 it holds that  $x_f \xrightarrow{\sigma}$  in  $\det(G)$  and  $I(z) = y$  and  $E(z) = \sigma$ . Since  $\det(G) \approx H_b$  it holds that there exists  $\tilde{x}_f \xrightarrow{\sigma}$  in  $H_b$ , where  $\tilde{x}_f \approx x_f$ . This means  $\tilde{y} \xrightarrow{\sigma} \tilde{z}$  in  $T'$ , where  $\tilde{y} = (\tilde{x}_d, \tilde{x}_f)$  and  $I(\tilde{z}) = (\tilde{x}_d, \tilde{x}_f)$  and  $E(\tilde{z}) = \sigma$ .
- $q_n = z$  is a  $Z$  state and  $p_n = z'$  is a  $Z$  state, Def. 12 (ii). Let  $z = ((x_d, x_f), e_o)$  and  $\sigma \in \Theta$ . Then based on Def. 12 it holds that  $\sigma \in \Sigma$  and  $I(z') = (x'_d, x_f)$  and  $x_d \xrightarrow{\sigma} x'_d$  in  $\det_d(G)$  and  $E(z') = e_o$ . Since  $\det_d(G) \approx H_{obd}$  and from  $x_d \xrightarrow{\sigma} x'_d$  in  $\det_d(G)$  it follows that  $\tilde{x}_d \xrightarrow{\sigma} \tilde{x}'_d$  in  $H_{obd}$ , where  $\tilde{x}_d \approx x_d$  and  $\tilde{x}'_d \approx x'_d$ , and based on the inductive assumption it follows that there exists  $\tilde{z} = ((\tilde{x}_d, \tilde{x}_f), e_o)$  such that  $\tilde{y}^0 \xrightarrow{\omega} \tilde{z}$  in  $T'$ . Thus,  $\tilde{z} \xrightarrow{\sigma} \tilde{z}'$  in  $T'$ , where  $\tilde{z} = ((\tilde{x}_d, \tilde{x}_f), e_o)$  and  $\sigma \in \Sigma$  and  $I(\tilde{z}') = (\tilde{x}'_d, \tilde{x}_f)$  and  $\tilde{x}_d \xrightarrow{\sigma} \tilde{x}'_d$  in  $H_{obd}$  and  $E(\tilde{z}') = e_o$ .
- $\tilde{q}_n \xrightarrow{\sigma} \tilde{p}_n$  is the  $\varepsilon$  insertion function and  $q_n = z$  is a  $Z$  state and  $p_n = w$  is a  $W$  state, Def. 12 (iii). Let  $z = ((x_d, x_f), e_o)$  and  $\sigma \in \Theta$ . Then based on Def. 12 it holds that  $\sigma = \varepsilon$  and  $I(w) = I(z)$  and  $A(w) = e_o$  and  $x_d \xrightarrow{e_a}$  and  $x_f \xrightarrow{e_a}$  in  $\det_d(G)$  and  $\det(G)$ , respectively. Since  $\det(G) \approx H_b$  and  $\det_d(G) \approx H_{obd}$ , from  $x_f \xrightarrow{e_a}$  in  $\det(G)$  it follows that  $\tilde{x}_f \xrightarrow{e_a}$  in  $H_b$  and from  $x_d \xrightarrow{e_a}$  in  $\det_d(G)$  it follows that  $\tilde{x}_d \xrightarrow{e_a}$  in  $H_{obd}$ . Moreover, based on the inductive assumption it holds that there exists  $\tilde{z} = ((\tilde{x}_d, \tilde{x}_f), e_o)$  such that  $\tilde{y}^0 \xrightarrow{\omega} \tilde{z}$

in  $T'$ . Thus,  $\tilde{z} \xrightarrow{\sigma} \tilde{w}$  in  $T'$ , where based on Def. 12 it holds that  $\sigma = \varepsilon$  and  $I(\tilde{w}) = I(\tilde{z})$  and  $A(\tilde{w}) = e_o$  and  $\tilde{x}_d \xrightarrow{e_a}$  and  $\tilde{x}_f \xrightarrow{e_a}$  in  $H_{obd}$  and  $H_b$ , respectively.

- $\tilde{q}_n \xrightarrow{\sigma} \tilde{p}_n$  is the event erasure transition function and  $q_n = z$  is a  $Z$  state and  $p_n = w$  is a  $W$  state, Def. 12 (iv). Let  $z = ((x_d, x_f), e_o)$  and  $\sigma \in \Theta$ . Then based on Def. 12 it holds that  $\sigma = e_o \rightarrow \varepsilon$  and  $I(w) = I(z)$  and  $A(w) = e_o \rightarrow \varepsilon$  and  $x_f \xrightarrow{e_a}$  in  $\det(G)$ . Since  $\det(G) \approx H_b$  from  $x_f \xrightarrow{e_a}$  in  $\det(G)$  it follows that  $\tilde{x}_f \xrightarrow{e_a}$  in  $H_b$ . Moreover, based on the inductive assumption it follows that there exists  $\tilde{z} = ((\tilde{x}_d, \tilde{x}_f), e_o)$  such that  $\tilde{y}^0 \xrightarrow{\omega} \tilde{z}$  in  $T'$ . Thus,  $\tilde{z} \xrightarrow{\sigma} \tilde{w}$  in  $T'$ , where based on Def. 12 it holds that  $\sigma = e_o \rightarrow \varepsilon$  and  $I(\tilde{w}) = I(\tilde{z})$  and  $A(\tilde{w}) = e_o \rightarrow \varepsilon$  and  $\tilde{x}_f \xrightarrow{e_a}$  in  $H_b$ .
- $q_n = w$  is a  $W$  state and  $p_n = y$  is a  $Y$  state and  $\sigma \in \Sigma$ , Def. 12 (v). Let  $w = ((x_d, x_f), e_o)$ . Then based on Def. 12 it holds that  $y = (x'_d, x'_f)$  and  $x_f \xrightarrow{e_a} x'_f$  in  $\det(G)$  and  $x_d \xrightarrow{e_a} x'_d$  in  $\det_d(G)$ . Since  $\det(G) \approx H_b$  and  $\det_d(G) \approx H_{obd}$ , from  $x_f \xrightarrow{e_a} x'_f$  in  $\det(G)$  it follows that  $\tilde{x}_f \xrightarrow{e_a} \tilde{x}'_f$  in  $H_b$  and from  $x_d \xrightarrow{e_a} x'_d$  in  $\det_d(G)$  it follows that  $\tilde{x}_d \xrightarrow{e_a} \tilde{x}'_d$  in  $H_{obd}$ . Moreover, based on the inductive assumption it holds that there exists  $\tilde{w} = ((\tilde{x}_d, \tilde{x}_f), e_o)$  such that  $\tilde{y}^0 \xrightarrow{\omega} \tilde{w}$  in  $T'$ . Thus,  $\tilde{w} \xrightarrow{\sigma} \tilde{y}$  in  $T'$ , where based on Def. 12 it holds that  $\tilde{y} = (\tilde{x}'_d, \tilde{x}'_f)$  and  $\tilde{x}_f \xrightarrow{e_a} \tilde{x}'_f$  in  $H_b$  and  $\tilde{x}_d \xrightarrow{e_a} \tilde{x}'_d$  in  $H_{obd}$ .
- $q_n = w$  is a  $W$  state and  $p_n = y$  is a  $Y$  state and  $\sigma \in \Sigma$ , Def. 12 (vi). Let  $w = ((x_d, x_f), e_o \rightarrow \varepsilon)$ . Then based on Def. 12 it holds that  $y = (x_d, x'_f)$  and  $x_f \xrightarrow{e_a} x'_f$  in  $\det(G)$ . Since  $\det(G) \approx H_b$  and from  $x_f \xrightarrow{e_a} x'_f$  in  $\det(G)$  it holds that  $\tilde{x}_f \xrightarrow{e_a} \tilde{x}'_f$  in  $H_b$ . Moreover, based on the inductive assumption it holds that there exists  $\tilde{w} = ((\tilde{x}_d, \tilde{x}_f), e_o \rightarrow \varepsilon)$  such that  $\tilde{y}^0 \xrightarrow{\omega} \tilde{w}$  in  $T'$ . Thus,  $\tilde{w} \xrightarrow{\sigma} \tilde{y}$  in  $T'$ , where based on Def. 12 it holds that  $\tilde{y} = (\tilde{x}_d, \tilde{x}'_f)$  and  $\tilde{x}_f \xrightarrow{e_a} \tilde{x}'_f$  in  $H_b$ .

( $\Leftarrow$ ) It must be shown that if  $\tilde{q}_n \xrightarrow{\sigma} \tilde{p}_n$  in  $T'$  then  $q_n \xrightarrow{\sigma} p_n$  in  $T$ . The same argument as ( $\Rightarrow$ ) holds. ■

**Theorem 8:** Let  $G_1 = \langle Q_1, \Sigma_1, \rightarrow_1, Q_1^0 \rangle$  and  $G_2 = \langle Q_2, \Sigma_2, \rightarrow_2, Q_2^0 \rangle$  be two nondeterministic automata with sets of secret states  $Q_1^S$  and  $Q_2^S$ , respectively. Let  $T_1 = \langle Q_1^1, Q_1^2, Q_1^3, \Sigma_1, \Sigma_1^e, \Theta_1, \rightarrow_{1, \Sigma_1^e}^1, \rightarrow_{1, \Sigma_1^e}^2, \rightarrow_{1, \Sigma_1^e}^3, y_1^0 \rangle$  and  $T_2 = \langle Q_2^1, Q_2^2, Q_2^3, \Sigma_2, \Sigma_2^e, \Theta_2, \rightarrow_{2, \Sigma_2^e}^1, \rightarrow_{2, \Sigma_2^e}^2, \rightarrow_{2, \Sigma_2^e}^3, y_2^0 \rangle$  be the largest three-player observers w.r.t.  $G_1$  and  $G_2$ , respectively. Let  $G_1^T = \langle \Sigma_{G_1^T}, Q_1, \rightarrow_1^T, Q_1^0 \rangle$  and  $G_2^T = \langle \Sigma_{G_2^T}, Q_2, \rightarrow_2^T, Q_2^0 \rangle$  be the transformed automata of  $T_1$  and  $T_2$ , respectively. Let  $T$  be the largest monolithic three-player observer w.r.t.  $G_1 \parallel G_2$ . Then let  $\rho : (\Sigma_{G_1^T} \cup \Sigma_{G_2^T}) \rightarrow (\Sigma_1 \cup \Sigma_1^e \cup \Theta_1) \cup (\Sigma_2 \cup \Sigma_2^e \cup \Theta_2)$  be a renaming. We have  $[G_1^T \parallel G_2^T \xrightarrow{s} (q_1, q_2)] \Rightarrow [T \xrightarrow{\rho(s)} q]$ .

*Proof:* We need to show that a transition is defined in  $G_1^T \parallel G_2^T$  if the equivalent transition is defined in  $T$ . It is shown by induction on  $n \geq 0$  that  $(y_1^0, y_2^0) \xrightarrow{s} (q_1^n, q_2^n)$  in  $G_1^T \parallel G_2^T$  implies  $y_T^0 \xrightarrow{\rho(s)} q_T$  in  $T$ .

Let  $G_1^T \parallel G_2^T \xrightarrow{s} (q_1, q_2)$ .

*Base case:*  $n = 0$ . Let  $(y_1^0, y_2^0)$  be the initial state of  $G_1 \parallel G_2$ , i.e.,  $y_1^0$  (respectively  $y_2^0$ ) is the initial state of  $T_1$  (respectively  $T_2$ ). From Def. 12,  $y_1^0 = (X_{1,obsd}^0, X_{1,obs}^0)$  and  $y_2^0 = (X_{2,obsd}^0, X_{2,obs}^0)$ , where  $X_{i,obs}^0$  and  $X_{i,obsd}^0$  are the initial state of  $\det(G_i)$  and  $\det_d(G_i)$  for  $i \in \{1, 2\}$ , respec-



tively. From Lemmas 6 and 7,  $\det(G_1 \parallel G_2)$  and  $\det_d(G_1 \parallel G_2)$  are isomorphic to  $\det(G_1) \parallel \det(G_2)$  and  $\det_d(G_1) \parallel \det_d(G_2)$ , which implies  $(X_{1,obs}^0, X_{2,obs}^0)$  is the initial state of  $\det(G_1) \parallel \det(G_2)$  and  $\det_d(G_1) \parallel \det_d(G_2)$ . Thus  $y_T^0 = ((X_{1,obsd}^0, X_{2,obsd}^0), (X_{1,obs}^0, X_{2,obs}^0))$  is the initial state of  $T$ .

*Inductive step:* Assume the claim holds for some  $0 \leq n$ , which means if  $(y_1^0, y_2^0) = (q_1^0, q_2^0) \xrightarrow{\sigma_0 \dots \sigma_{n-1}} (q_1^n, q_2^n)$  in  $G_1^T \parallel G_2^T$  then  $y^0 \xrightarrow{\rho(\sigma_0 \dots \sigma_{n-1})} q^n$  in  $T$ . Now we need to show that if  $(q_1, q_2) = (q_1^n, q_2^n) \xrightarrow{\sigma_n} (p_1, p_2)$  in  $G_1^T \parallel G_2^T$  then  $q = q^n \xrightarrow{\rho(\sigma_n)} p$  in  $T$ . From  $(q_1, q_2) = (q_1^n, q_2^n) \xrightarrow{\sigma_n} (p_1, p_2)$  in  $G_1^T \parallel G_2^T$  and based on Def. 2 it holds that  $q_i \xrightarrow{\sigma_n} p_i$  in  $G_i^T$  for  $i \in \{1, 2\}$ , which means  $q_i \xrightarrow{\sigma_n} p_i$  in  $T_i$  for  $i \in \{1, 2\}$ . Consider the following four cases for all the possible transitions:

- if  $\rho(\sigma_n) = \sigma_n$  then based on Def. 24 it holds that  $q_i \xrightarrow{\sigma_n} p_i$  is a yz transition in the original  $T_i$  such that  $E(p_i) = \sigma_n$  and  $I(p_i) = q_i$  if  $\sigma_n \in \Sigma_i$  and  $q_i = p_i$  otherwise for  $i \in \{1, 2\}$ . Let  $q_i = (x_{i,d}, x_{i,f})$  for  $i \in \{1, 2\}$ . Based on Def. 12 this means  $x_{i,f} \xrightarrow{\sigma_n} p_{i,f}$  in  $\det(G_i)$  if  $\sigma_n \in \Sigma_i$ . Moreover, based on the inductive assumption there exists  $y = (x_d, x_f)$  such that  $y_T^0 \xrightarrow{\omega} y$  in  $T$ , which implies  $\det(G_1 \parallel G_2) \xrightarrow{P_e(\omega)} x_f$ . Since based on Lemma 6  $\det(G_1 \parallel G_2)$  and  $\det(G_1) \parallel \det(G_2)$  are isomorphic it holds  $\det(G_1) \parallel \det(G_2) \xrightarrow{P_e(\omega)} x_f$  and string  $P_e(\omega)$  reaches  $(x_{1,f}, x_{2,f})$  in  $\det(G_1) \parallel \det(G_2)$ . Thus, based on  $x_{i,f} \xrightarrow{\sigma_n} p_{i,f}$  in  $\det(G_i)$  it can be deduced that  $(x_{1,f}, x_{2,f}) \xrightarrow{\sigma_n} (p_{1,f}, p_{2,f})$  in  $\det(G_1 \parallel G_2)$ , it also implies  $(x_{1,f}, x_{2,f}) \xrightarrow{\sigma_n} (p_{1,f}, p_{2,f})$  in  $\det(G_1 \parallel G_2)$  by Lemma 6. This means  $q_T^n = q_T \xrightarrow{\sigma_n} p_T$  is a yz transition in  $T$ .
- if  $\rho(\sigma_n) = \alpha$  and  $\sigma_n = \alpha_{e_o}$ . Then based on Def. 24 there are three possibilities for  $q_i \xrightarrow{\rho(\sigma_n)} p_i$  in  $T_i$ : it is a zz transition or a zw<sub>1</sub> transition or a zw<sub>2</sub> transition and  $e_o = E(q_i)$ . Now consider the following cases:

- 1)  $\alpha_{e_o} \in \Sigma_{G_1^T} \setminus \Sigma_{G_2^T}$ . This means  $q_1 \xrightarrow{\alpha_{e_o}} p_1$  in  $G_1^T$ , which implies  $q_1 \xrightarrow{\sigma_n} p_1$  in  $T_1$ . From  $\sigma_{ne_o} \notin \Sigma_{G_2^T}$  and def. 24 it follows that  $\sigma_n \notin \Sigma_2$ , which implies  $q_2 = p_2$ . Consider the following three cases: If  $q_1 \xrightarrow{\sigma_n} p_1$  in  $T_1$  is a zz transition, then based on Def. 12 it holds that  $x_{1,d} \xrightarrow{\sigma_n} x'_{1,d}$  in  $\det_d(G_1)$ . If  $q_1 \xrightarrow{\sigma_n} p_1$  in  $T_1$  is a zw<sub>1</sub> transition, then based on Def. 12 it holds that  $x_{1,d} \xrightarrow{\sigma_n}$  in  $\det_d(G_1)$  and  $x_{1,f} \xrightarrow{\sigma_n}$  in  $\det(G_1)$ . If  $q_1 \xrightarrow{\sigma_n} p_1$  in  $T_1$  is a zw<sub>2</sub> transition, then based on Def. 12 it holds that  $x_{1,f} \xrightarrow{\sigma_n}$  in  $\det(G_1)$ . In all the three cases based on Lemmas 6 and 7, where they show  $\det_d(G_1 \parallel G_2)$  and  $\det_d(G_1) \parallel \det_d(G_2)$  are isomorphic, it holds that  $(x_{1,d}, x_{2,d}) \xrightarrow{\sigma_n} (x'_{1,d}, x_{2,d})$  in  $\det_d(G_1 \parallel G_2)$  and  $(x_{1,f}, x_{2,f}) \xrightarrow{\sigma_n}$  in  $\det(G_1 \parallel G_2)$ . These mean  $q^n = q \xrightarrow{\sigma_n} p$  in  $T$  is a zz transition if  $q_1 \xrightarrow{\sigma_n} p_1$  in  $T_1$  is a zz transition,  $q^n = q \xrightarrow{\sigma_n} p$  is a zw<sub>1</sub> transition in  $T$  if  $q_1 \xrightarrow{\sigma_n} p_1$  is a zw<sub>1</sub> transition in  $T_1$  and  $q_T^n = q_T \xrightarrow{\sigma_n} p_T$  is a zw<sub>2</sub> transition in  $T$  if  $q_1 \xrightarrow{\sigma_n} p_1$  is a zw<sub>2</sub> transition in  $T_1$ .
- 2)  $\sigma_{ne_o} \in \Sigma_{G_1^T} \cap \Sigma_{G_2^T}$ . This means  $q_i \xrightarrow{\sigma_{ne_o}} p_i$  in  $G_i^T$  for  $i \in \{1, 2\}$ , which implies  $q_i \xrightarrow{\sigma_n} p_i$  in  $T_i$  for  $i \in \{1, 2\}$ .

Again there are three cases: If  $q_i \xrightarrow{\sigma_n} p_i$  in  $T_i$  for  $i = 1, 2$  is a zz transition. Then based on Def. 12 it holds that  $x_{i,d} \xrightarrow{\sigma_n} x'_{i,d}$  in  $\det_d(G_i)$  for  $i = 1, 2$ . If  $q_i \xrightarrow{\sigma_n} p_i$  in  $T_i$  is a zw<sub>1</sub> transition for  $i \in \{1, 2\}$ . Then based on Def. 12 it holds that  $x_{i,d} \xrightarrow{\sigma_n}$  in  $\det_d(G_i)$  and  $x_{i,f} \xrightarrow{\sigma_n}$  in  $\det(G_i)$  for  $i = 1, 2$ . If  $q_i \xrightarrow{\sigma_n} p_i$  in  $T_i$  is a zw<sub>2</sub> transition for  $i = 1, 2$ . Then based on Def. 12 it holds that  $x_{i,f} \xrightarrow{\sigma_n}$  in  $\det(G_i)$  for  $i = 1, 2$ . In all the three cases based on Lemmas 6 and 7 where they show  $\det_d(G_1 \parallel G_2)$  and  $\det_d(G_1) \parallel \det_d(G_2)$  are isomorphic, it holds that  $(x_{1,d}, x_{2,d}) \xrightarrow{\sigma_n} (x'_{1,d}, x'_{2,d})$  in  $\det_d(G_1 \parallel G_2)$  and  $(x_{1,f}, x_{2,f}) \xrightarrow{\sigma_n}$  in  $\det(G_1 \parallel G_2)$ . These mean  $q_T^n = q_T \xrightarrow{\sigma_n} p_T$  in  $T$  is a zz transition if  $q_i \xrightarrow{\sigma_n} p_i$  in  $T_i$  is a zz transition,  $q_T^n = q_T \xrightarrow{\sigma_n} p_T$  is a zw<sub>1</sub> transition in  $T$  if  $q_i \xrightarrow{\sigma_n} p_i$  is a zw<sub>1</sub> transition in  $T_i$  and  $q_T^n = q_T \xrightarrow{\sigma_n} p_T$  is a zw<sub>2</sub> transition in  $T$  if  $q_i \xrightarrow{\sigma_n} p_i$  is a zw<sub>2</sub> transition in  $T_i$  for  $i \in \{1, 2\}$ .

3)  $\sigma_{ne_o} \in \Sigma_{G_2^T} \setminus \Sigma_{G_1^T}$ . The same argument as case 1.

- if  $\rho^{-1}(\sigma_n) = \sigma_{ne_o, w}$  then based on Def. 24 there are again three cases:

- 1)  $\sigma_{ne_o, w} \in \Sigma_{G_1^T} \setminus \Sigma_{G_2^T}$ . This means  $q_1 \xrightarrow{\sigma_{ne_o, w}} p_1$  in  $G_1^T$ . From  $\sigma_{ne_o, w} \notin \Sigma_{G_2^T}$  and def. 24 it follows that  $\sigma_n \notin \Sigma_2$ , which implies  $q_2 = p_2$ . From  $\rho^{-1}(\sigma_n) = \sigma_{ne_o, w}$  it holds that  $p_1 \xrightarrow{\sigma_n} q_1$  in  $T_1$  is a wyl transition. This means  $x_{1,d} \xrightarrow{\sigma_n} x'_{1,d}$  in  $\det_d(G_1)$  and  $x_{1,f} \xrightarrow{\sigma_n} x'_{1,f}$  in  $\det(G_1)$ . Based on Lemmas 6 and 7, where they show  $\det_d(G_1 \parallel G_2)$  and  $\det_d(G_1) \parallel \det_d(G_2)$  are isomorphic, it holds that  $(x_{1,d}, x_{2,d}) \xrightarrow{\sigma_n} (x'_{1,d}, x_{2,d})$  in  $\det_d(G_1 \parallel G_2)$  and  $(x_{1,f}, x_{2,f}) \xrightarrow{\sigma_n} (x'_{1,f}, x_{2,f})$  in  $\det(G_1 \parallel G_2)$ . This means  $q_T^n = q_T \xrightarrow{\sigma_n} p_T$  is a wyl transition in  $T$ .
- 2)  $\sigma_{ne_o, w} \in \Sigma_{G_1^T} \cap \Sigma_{G_2^T}$ . Then it follows that  $q_i \xrightarrow{\sigma_{ne_o, w}} p_i$  in  $G_i^T$  for  $i \in \{1, 2\}$ . This means  $q_i \xrightarrow{\sigma_n} p_i$  in  $T_i$  is a wyl transition, which implies  $x_{i,d} \xrightarrow{\sigma_n} x'_{i,d}$  in  $\det_d(G_i)$  and  $x_{i,f} \xrightarrow{\sigma_n} x'_{i,f}$  in  $\det(G_i)$  for  $i \in \{1, 2\}$ . Based on Lemmas 6 and 7, where they show  $\det_d(G_1 \parallel G_2)$  and  $\det_d(G_1) \parallel \det_d(G_2)$  are isomorphic, it holds that  $(x_{1,d}, x_{2,d}) \xrightarrow{\sigma_n} (x'_{1,d}, x'_{2,d})$  in  $\det_d(G_1 \parallel G_2)$  and  $(x_{1,f}, x_{2,f}) \xrightarrow{\sigma_n} (x'_{1,f}, x'_{2,f})$  in  $\det(G_1 \parallel G_2)$ . This means  $q_T^n = q_T \xrightarrow{\sigma_n} p_T$  is a wyl transition in  $T$ .
- 3)  $\sigma_{ne_o, w} \in \Sigma_{G_2^T} \setminus \Sigma_{G_1^T}$ . The same argument as case 1.

- if  $\rho^{-1}(\sigma_n) = \sigma_{ne_o \rightarrow e, w}$  then the same argument as above,  $\rho^{-1}(\sigma_n) = \sigma_{ne_o, w}$ , holds. ■

**Theorem 11:** Let  $G = \langle \Sigma, Q, \rightarrow, Q^0 \rangle$  be a nondeterministic automaton with the set of secret states  $Q^S \subseteq Q$  and let  $\Phi$  be the edit constraint which prohibits  $n+1$  consecutive event erasures. Let  $T = \langle Q_Y, Q_Z, Q_W, \Sigma, \Sigma^E, \Theta, \rightarrow_{yz}, \rightarrow_{zz}, \rightarrow_{zw}, \rightarrow_{wy}, y_0 \rangle$  be the largest three-player observer w.r.t.  $G$  without considering the edit constraint and let  $G^T = \langle \Sigma_{G^T}, Q, \rightarrow, Q^0 \rangle$  be the transformed automaton of  $T$ . Let  $\rho : \Sigma_{G^T} \rightarrow (\Sigma_1 \cup \Sigma_1^E \cup \Theta_1)$  be a renaming and let AES be the All Edit Structure obtained from  $T$ . Let  $S$  be the supremal controllable and nonblocking subautomaton of  $G^T$  after considering the specification intro-

duced by  $\Phi$ . Then  $S \xrightarrow{\rho^{-1}(s)} q$  if and only if  $AES \xrightarrow{\rho(s)} q$ .

*Proof:* The pruning process to obtain the AES and the supervisor synthesis procedure are both iterative, which remove states at each iteration. We will show by induction that at each iteration a state is removed in  $G^T$  by the supervisory control synthesis procedure if and only if the corresponding state is removed by the pruning process from  $T$ .

*Base case:* Clearly  $G^T$  and  $T$  have the same transition relation.

*Inductive step:* Assume the claim holds for some  $n > 0$ . We let  $X_G^n$  be the state space of  $G^T$  at the  $n$ -th iteration of supervisor synthesis process and  $X_T^n$  be the state set of  $T$  at the  $n$ -th iteration of pruning process. Then  $S \xrightarrow{\rho^{-1}(s)} q \Leftrightarrow AES \xrightarrow{\rho(s)} q$  holds for all  $q \in X_G^n$  ( $q \in X_T^n$ ). Now we need to show that  $X_G^{n+1}$  and  $X_T^{n+1}$  are also equal. Assume  $G^T \xrightarrow{s} q$ , which implies  $T \xrightarrow{\rho(s)} q$  based on Def. 24.

( $\Rightarrow$ ) First we show that if  $q \notin X_T^{n+1}$ , which means  $q$  is removed by the pruning process at the  $n$ -th iteration, so  $q \notin X_G^{n+1}$ . Then  $q \notin X_T^{n+1}$  if it is a deadlock state or it is a  $Y$  state and there exists  $e_o \in \Sigma$  such that  $q \xrightarrow{e_o} z'$ , where  $z'$  is a deadlock  $Z$  state. If  $q$  is a deadlock state then based on Def. 12 it holds that  $q$  is either a  $W$  state or a  $Z$  state. Thus, consider the following three case:

- $q$  is a  $W$  state. Then  $\exists e_o \in \Sigma$  such that  $q \xrightarrow{e_o}$ . Then based on Def. 24 it holds that either  $q \not\xrightarrow{e_o} \epsilon$  or  $q \xrightarrow{e_o} \epsilon$  in  $G^T$  either, which means  $q \not\vdash$ . Thus,  $q$  is a blocking state in  $G^T$  and  $q \notin \Theta^{nonb}(X_G^n)$ .
- $q$  is a  $Z$  state. If  $q$  is a deadlock state then  $\exists \theta \in \Theta$  such that  $q \xrightarrow{\theta} z'$  or  $q \xrightarrow{\theta} w$  in  $T$ . Then based on Def. 24 it holds that either  $q \xrightarrow{\theta_{E(q)}} z'$  or  $q \xrightarrow{\theta_{E(q)}} w$  does not exist in  $G^T$ , which means  $q \not\vdash$ . Thus,  $q$  is a blocking state in  $G^T$  and  $q \notin \Theta^{nonb}(X_G^n)$ .
- $q$  is a  $Y$  state and  $q \xrightarrow{e_o} z'$ , where  $z'$  is a deadlock  $Z$  state. Then based on Def. 24 it holds that  $q \xrightarrow{e_o} z'$  in  $G^T$  and  $e_o \in \Sigma_u$ . As it was shown above if  $z'$  is a deadlock state in  $T$  then  $z'$  is also a deadlock state in  $G^T$ , thus removed by the supervisor synthesis procedure. If  $z'$  is removed, i.e.,  $z' \notin X_G^n$ , then  $q \notin \Theta^{cont}(X_G^n)$ .

Thus, if  $q$  is removed in the pruning of  $T$ , then  $q$  is also removed from  $G^T$  in supervisor synthesis.

( $\Leftarrow$ ) Now we show that if  $q \notin \Theta^{nonb}(X_G^n) \cap \Theta^{cont}(X_G^n)$ , then  $q \notin X_T^{n+1}$  and  $q$  needs to be removed from  $T$  by the pruning process. If  $q \notin \Theta^{nonb}(X_G^n) \cap \Theta^{cont}(X_G^n)$ , there are two cases:

- $q \notin \Theta^{nonb}(X_G^n)$ . Then it holds that  $q$  is a blocking state, which means  $q \not\vdash$  does not exists in  $G^T$ . There are three possibilities for  $q$ , it can be a  $Y$ ,  $Z$  or  $W$  state. If  $q$  is a  $Y$  state in  $T$  then  $q \in Q^m$  in  $G$ , which means  $q \in \Theta^{nonb}$ , which contradicts the assumption. Thus,  $q$  can only be a  $W$  or  $Z$  state in  $T$ . In both cases from  $q \not\vdash$  in  $G^T$  and Def. 24, it holds that  $\exists \rho(\alpha) \in \Theta$  such that  $q \xrightarrow{\rho(\theta)}$  in  $T$ . This means  $q$  is a deadlock state in  $T$  and  $q \notin X_T^{n+1}$ .
- $q \notin \Theta^{cont}(X_G^n)$ . This means that  $q \xrightarrow{u} z$  in  $G^T$  such that  $u \in \Sigma_u$  and  $z \notin X_G^n$ . Based on Def. 24, this means that  $q$  is a  $Y$  state in  $T$  and  $z$  is a  $Z$  state. It was shown above that if  $z \notin X_G^n$  then  $z \notin X_T^n$ , which means  $z$  is a deadlock  $Z$

state. Thus  $q$  will be removed by pruning process, which means  $q \notin X_T^{n+1}$ .

Thus, if  $q$  is removed by synthesis from  $G^T$  then  $q$  is also removed from  $T$  by pruning. ■