Fakultät für Elektrotechnik und Informationstechnik
Technische Universität München

Secure-by-Construction Synthesis of Cyber-Physical Systems

Siyuan Liu

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades einer

Doktorin der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender:
Prof. Dr.-Ing. Wolfgang Kellerer

Prüfende der Dissertation:
1. Prof. Dr.-Ing./Univ. Tokio Martin Buss
2. Prof. Dr. Majid Zamani
3. Prof. Dr. Dimos V. Dimarogonas

Die Dissertation wurde am 12.04.2022 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 21.06.2022 angenommen.

This dissertation is dedicated to my parents and my better half, for their unconditional love and endless support.

Acknowledgments

This dissertation summarizes the result over the last years during my doctoral research studies in the Department of Electrical and Computer Engineering at the Technical University of Munich (TUM), Germany. Here, I would like to take the opportunity to acknowledge those people who supported me to make the completion of this thesis possible.

First and foremost, I would like to express my deepest gratitude to my advisor Prof. Majid Zamani who guided me through my Ph.D. journey with his wisdom, enthusiasm, and patience. I sincerely appreciate his encouragement to pursue an academic career. His continuous guidance, insightful vision and generous support have made this thesis possible and will inspire me to be a better researcher in my future life.

I wish to sincerely extend my gratitude to Prof. Martin Buss for welcoming me in his research group in the Chair of Automatic Control Engineering at the Technical University of Munich, Germany, since July 2019. I would greatly appreciate all his generous help, support, and consideration during this time.

Besides, my deep thanks go to Prof. Dimos V. Dimarogonas for providing me with a great opportunity to visit his group at the KTH Royal Institute of Technology, Stockholm, Sweden, and for the many fruitful discussions we had that broadened my view on research. I am sincerely honored to have him as part of my thesis committee. My thanks also go to his Ph.D. students and the friends I met at the Division of Decision and Control Systems for the welcoming atmosphere and worthwhile discussions.

I want to express my thanks to Prof. Xiang Yin at Shanghai Jiao Tong University for his helpful and inspiring instructions during the many collaborations we had in the past years and for being a role model for me towards becoming a mature researcher.

I would like to thank all my colleagues at HyConSys Lab for creating a warm and friendly working atmosphere, and for the inspiring discussions and enjoyable times we shared during the last five years. I am fortunate to be a member of HyConSys Lab and privileged to meet them all in person in Munich and virtually in Boulder.

No words can express my gratitude to my family. I would not have been able to even start my Ph.D. journey without the encouragement of my family and the love of my better half. I am genuinely grateful to my parents for always believing in me. I am deeply indebted to my mother for her endless support and unconditional love. Last but not least, I wish to express my special thanks to my husband Dalong, who has always been there for me along this journey, for his tender love, invaluable patience, and endless support.

Siyuan Liu
Munich, April 2022

Abstract

Cyber-physical systems (CPS) are the technological backbone of the increasingly interconnected and smart world where design faults or security vulnerability can be catastrophic. Implantable medical devices, smart buildings, and critical infrastructure are high-profile examples that underscore security and safety concerns of modern CPS. In the last decade, safety concerns received considerable attention in the design of CPS, while security analysis is left as an afterthought for later stages. This paradigm results in costly and lengthy development processes due to high validation costs. This thesis advocates a paradigm shift in the development of CPS by proposing a secure-by-construction synthesis scheme that generalizes existing correct-by-construction synthesis methods by considering security properties in addition to safety ones during the design phase. Our focus is to develop theoretical foundations to bridge the gap between control theory and theoretical computer science on the analysis of security properties.

The first step to bridging the gap is to provide a common framework that generalizes the security notions from different research fields of discrete-event systems (DES), control theory, and formal methods. We develop a generalized security notion by integrating the ideas from distinct research areas. A new notion of approximate opacity is proposed that is more applicable to CPS by quantitatively evaluating the security level concerning the measurement precision of malicious intruders.

Two main approaches are presented in this thesis to analyze opacity for complex CPS. The first approach provides abstraction-based frameworks to verify approximate opacity for both stochastic and non-stochastic systems. By introducing new notions of opacity-preserving simulation relations, we construct opacity-preserving finite abstractions mimicking the behaviors of complex CPS, which enable us to verify opacity of concrete (stochastic) CPS by using their finite abstractions. The second approach provides an alternative discretization-free strategy for verifying opacity via a notion of barrier certificates. This approach is a deductive method that provides sufficient conditions for approximate opacity of complex CPS. Two notions of barrier certificates are proposed which are used in reverse directions in the sense that one guarantees opacity and the other ensures the lack of opacity of the system.

To overcome the challenges encountered with large-scale CPS, we further develop modular approaches to reduce the computational complexity of the proposed opacity verification schemes. By breaking the large-scale system into semi-independent parts, we show that the verification problem can be addressed in a cost-efficient way by assuming some small-gain type conditions. Compositional construction of opacity-preserving abstractions is developed for both interconnected control systems and switched systems. A compositionality result is further derived to compute barrier certificates for verifying opacity of interconnected control systems.

Zusammenfassung

Cyber-physische Systeme (CPS) sind das technologische Rückgrat der zunehmend vernetzten und intelligenten Welt, in der Konstruktionsfehler oder Sicherheitslücken katastrophale Folgen haben können. Implantierbare medizinische Geräte, intelligente Gebäude und kritische Infrastrukturen sind prominente Beispiele, die die Sicherheitsbedenken moderner CPS unterstreichen. In den letzten zehn Jahren wurde den Sicherheitsaspekten bei der Entwicklung von CPS große Aufmerksamkeit geschenkt, während die Sicherheitsanalyse erst in späteren Phasen berücksichtigt wurde. Dieses Paradigma führt zu kostspieligen und langwierigen Entwicklungsprozessen, da die Validierungskosten hoch sind. Diese Arbeit befürwortet einen Paradigmenwechsel in der Entwicklung von CPS, indem sie ein Secure-by-Construction-Syntheseschema vorschlägt, das die bestehenden Correct-by-Construction-Synthesemethoden verallgemeinert Synthesemethoden verallgemeinert, indem zusätzlich zu den Sicherheitseigenschaften auch die Sicherheitseigenschaften während der Entwurfsphase berücksichtigt werden. Unser Ziel ist es, theoretische Grundlagen zu entwickeln, um die Lücke zwischen Kontrolltheorie und theoretischer Informatik bei der Analyse von Sicherheitseigenschaften zu schließen.

Der erste Schritt zur Überbrückung der Kluft besteht darin, einen gemeinsamen Rahmen zu schaffen, der die Sicherheitsbegriffe aus den verschiedenen Forschungsbereichen der ereignisdiskreten Systeme (DES), der Kontrolltheorie und der formalen Methoden verallgemeinert. Wir entwickeln ein verallgemeinertes Sicherheitskonzept, indem wir die Ideen aus verschiedenen Forschungsbereichen integrieren. Es wird ein neuer Begriff der ungefähren Undurchsichtigkeit vorgeschlagen, der besser auf CPS anwendbar ist, indem das Sicherheitsniveau in Bezug auf die Messgenauigkeit von böswilligen Eindringlingen quantitativ bewertet wird.

In dieser Arbeit werden zwei Hauptansätze zur Analyse der Opazität für komplexe CPS vorgestellt. Der erste Ansatz bietet einen auf Abstraktion basierenden Rahmen, um die ungefähre Opazität sowohl für stochastische als auch für nicht-stochastische Systeme zu überprüfen. Durch die Einführung neuer Begriffe für opazitätserhaltende Simulationsrelationen konstruieren wir opazitätserhaltende endliche Abstraktionen, die das Verhalten komplexer CPS nachahmen und uns in die Lage versetzen, die Opazität konkreter (stochastischer) CPS mit Hilfe ihrer endlichen Abstraktionen zu verifizieren. Der zweite Ansatz bietet eine alternative diskretisierungsfreie Strategie zur Verifizierung der Opazität über den Begriff der Barrierezertifikate. Dieser Ansatz ist eine deduktive Methode, die hinreichende Bedingungen für die ungefähre Opazität komplexer CPS liefert. Es werden zwei Begriffe für Barrierezertifikate vorgeschlagen, die in umgekehrter Richtung verwendet werden können, in dem Sinne, dass einer die Opazität garantiert und der andere das Fehlen der Opazität des Systems sicherstellt.

Um die Herausforderungen zu bewältigen, die bei großen CPS auftreten, entwickeln wir modulare Ansätze, um die Rechenkomplexität der vorgeschlagenen Verifizierungsverfahren zu reduzieren. Indem wir das große System in halb unabhängige Teile zerlegen, zeigen wir, dass das Verifizierungsproblem auf kosteneffiziente Weise angegangen werden kann, indem wir Bedingungen vom Typ small-gain annehmen. Es wird eine kompositionelle Konstruktion von opazitätserhaltenden Abstraktionen sowohl für vernetzte Kontrollsysteme als auch für geschaltete Systeme entwickelt. Darüber hinaus wird ein Kompositionalitätsergebnis abgeleitet, um Barrierezertifikate zu berechnen, die auf vernetzte Kontrollsysteme zugeschnitten sind.

Publications by the Author during Ph.D.

Under Submission and Review

1. J. Hou, **S. Liu**, X. Yin, and M. Zamani, "Abstraction-based verification of approximate pre-opacity for control systems," under review.
2. **S. Liu**, X. Yin, and M. Zamani, "On approximate opacity of stochastic control systems," under submission.

Journal Papers

3. **S. Liu**, A. Trivedi, X. Yin, and M. Zamani, "Secure-by-construction synthesis of cyber-physical systems," *Annual Reviews in Control*, vol. 53, pp. 30–50, 2022.
4. **S. Liu**, N. Noroozi, and M. Zamani, "Symbolic models for infinite networks of control systems: A compositional approach," *Nonlinear Analysis: Hybrid Systems*, vol. 43, 2021.
5. **S. Liu**, A. Swikir, and M. Zamani, "Verification of initial-state opacity for switched systems: A compositional approach," *Nonlinear Analysis: Hybrid Systems*, vol. 42, 2021.
6. S. Tasdighi Kalat, **S. Liu**, and M. Zamani, "Modular verification of opacity for interconnected control systems via barrier certificates," *IEEE Control Systems Letters*, vol. 6, pp. 890–895, 2021.
7. **S. Liu** and M. Zamani, "Compositional synthesis of opacity-preserving finite abstractions for interconnected systems," *Automatica*, vol. 131, 2021.
8. **S. Liu** and M. Zamani, "Verification of approximate opacity via barrier certificates," *IEEE Control Systems Letters*, vol. 5, no. 4, pp. 1369–1374, 2020.
9. X. Yin, M. Zamani, and **S. Liu**, "On approximate opacity of cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1630–1645, 2020.

Conference Papers

10. **S. Liu**, A. Saoud, P. Jagtap, D. V. Dimarogonas, and M. Zamani, "Compositional synthesis of signal temporal logic tasks via assume-guarantee contracts," *61st Conference on Decision and Control (CDC)*, to appear, December 2022.
11. S. Tasdighi Kalat, **S. Liu**, and M. Zamani, "Verification of approximate infinite-step opacity using barrier certificates," *European Control Conference (ECC)*, pp. 175–180, 2022.

12. **S. Liu**, A. Swikir, and M. Zamani, "Compositional verification of initial-state opacity for switched systems," *59th Conference on Decision and Control (CDC)*, pp. 2146–2151, December 2020.
13. **S. Liu**, X. Yin, and M. Zamani, "On a notion of approximate opacity for discrete-time stochastic control systems," *American Control Conference (ACC)*, pp. 5413–5418, July 2020.
14. **S. Liu** and M. Zamani, "Compositional synthesis of almost maximally permissible safety controllers," *American Control Conference (ACC)*, pp. 1678–1683, July 2019.

Contents

List of Figures

List of Abbreviations

ACBC	augmented control barrier certificate
ASR	alternating simulation relations
CEGIS	counter-example-guided inductive synthesis
CPS	cyber-physical system
DES	discrete-event system
CurSOPSF	current-state opacity-preserving simulation function
dt-CS	discrete-time control system
dt-SCS	discrete-time stochastic control system
dt-SS	discrete-time switched system
δ-ISS	incremental input-to-state stability
gMDP	general Markov decision process
InitSOPSF	initial-state opacity-preserving simulation function
InfSOPSF	infinite-step opacity-preserving simulation function
IoT	internet-of-things
i.i.d.	independent and identically distributed
LTL	linear temporal logic
MDP	Markov decision process
ODE	ordinary differential equation
OFRR	output-feedback refinement relation
RL	reinforcement learning
SCC	strongly connected component
SMT	satisfiability modulo theory
SOS	sum-of-squares
SSF	stochastic simulation functions

1 Introduction

1.1 Motivation

The revolution in miniaturized communication devices in the beginning of this millennium contributed towards a revolution in the internet-of-things (IoT) and the networked systems woven around them: the cyber-physical systems (CPS). CPS are marked by a close-knit interaction of discrete computation and continuous control over a network and are playing critical roles in virtually every aspect of our modern experience ranging from consumer electronics to implantable medical devices, from smart cars to smart hospitals, and from controlling our power systems to safeguarding our nuclear rectors. These systems are clearly safety-critical as a bug in their design could be life threatening, but given their societal implications, they are also security-critical where a bug in their design may have the potential to jeopardize the privacy, trust, and economic interests of society built around them. In the last decade, safety concerns have received considerable attention in the design of CPS, while security analysis is left as an afterthought for the later stages. This existing paradigm results in costly and lengthy development of CPS due to very high security verification and validation costs. We believe that the security considerations should be elevated as primary design drivers along with safety ones to tackle the design challenge of modern CPS and call for a need to expand the correct-by-construction paradigm of designing safety-critical systems to encompass security: we call this paradigm *secure-by-construction*.

Security considerations in the traditional computer science literature are often classified along the CIA mnemonic: *confidentiality*, *integrity*, and *availability*. The confidentiality properties concern the protection of sensitive information leakage either directly or, more importantly, via side-channels (seemingly harmless observations of the system by malintent eavesdroppers). The umbrella-term integrity targets the establishment of the trust in the authenticity of the source of the information. Finally, availability properties concern with the protection of the system operations from cyberattacks aimed at disrupting or interrupting the core functionality of the system. While ensuring integrity deals with similar issues as for classical computer systems and can benefit from current best practices on encryption, the confidentiality and availability concerns in CPS get amplified due to a plethora of attack surfaces available in the form of physical system observations and constraints ranging from the usual time and memory to temperature, acoustics, pressure, and electro-magnetic radiation.

On the positive side, since principled approaches to CPS modeling and analysis already embrace the integration of the encoding of physical variables and discrete control, the confidentiality and availability properties can be explicated during the design time to ensure a system that is not only functional, but also guarantees freedom from

known vulnerabilities. This is primary tenet of our stance on CPS-security: the design of security-critical CPS must tackle both functionality and security challenges simultaneously by leveraging correct-by-construction synthesis to include confidentiality and availability.

Security-related attacks are increasingly becoming pervasive in safety-critical CPS. While most of the well-known attacks—such as drone hacking [196], Jeep hacking [54], pacemaker and Implantable Cardioverter Defibrillator (ICD) attacks [62, 152]—exploit unencrypted wireless communication, such attacks can be readily guarded against by following recommended cryptographic measures without requiring any significant modification to the control logic. On the other hand, security vulnerabilities related to information leaks via side-channels may be impossible to mitigate without requiring a non-trivial modification to control software, as the side-channels are products of the interaction of the embedded control software with its physical environment.

To provide a simple scenario of unintended information leak via timing side-channels, let us consider an example in the setting of smart hospitals shown in Figure 1.1. An increasing prevalence of smart-devices and sensors in modern hospitals makes such an attack scenario on smart hospitals viable. While at a first glance, this example may seem contrived, it emphasizes how seemingly innocuous observations can provide a strong side-channel to leak private information. Furthermore, the presence of wide variety of observations (time delays between various responses [104], temperature [73], electro-magnetic emissions [124], optical [124] and acoustic [49], physiological [133]) in CPS expose corresponding attack surfaces to the intruder and render CPS even more vulnerable than traditional software.

Formal-methods based approach to system design [186, 17] recommends rigorous requirement specification in every stage of the system development. Formal verification [13] and controller synthesis [186, 17] are two leading approaches to provide correctness guarantees with respect to such requirements. While formal verification aims at providing a proof of correctness with respect to the given specifications, the goal of the controller synthesis approach is more ambitious: it takes a control system together with the specification, and produces a controller such that the resulting closed-loop satisfies the specification. The automated controller synthesis approach from formal requirements is referred to as correct-by-construction controller synthesis scheme [186, 17, 101]. While the controller synthesis approach has been well understood for safety, the security requirements in CPS are often verified after the design of controllers. Hence, if the system leaks information, the controller needs to be redesigned incurring high verification and validation costs.

We envisage a paradigm shift in the development of simultaneously safe and secure CPS that advocates a **secure-by-construction** synthesis scheme which generalizes existing correct-by-construction synthesis methods by considering privacy properties simultaneously to safety ones during the design phase.

Overview We give a brief overview of the secure-by-construction approach using a concrete synthesis problem for our experimental setup. Consider a physical platform

Services provided by average consultation, examination, and wait times

Service	**Avg. Total Time (min)**	**Avg. Total Wait (min)**	**Avg. Time with Nurse (min)**	**Avg. Time with Physician (min)**	**No. Cases (n)**	**Incomplete Cases (n)**
Minor assessment (std.)	50 (30)	33 (22)	1 (3)	16 (13)	67	11
Intermediate assessment (std.)	55 (24)	37 (21)	2 (3)	16 (12)	400	29
General assessment (std.)	77 (27)	31 (17)	10 (5)	36 (19)	43	1
Psychotherapy (std.)	71 (22)	35 (16)	2 (3)	34 (14)	11	0
Annual exam (after 16^{th} birthday) (std.)	51 (30)	26 (12)	7 (4)	18 (8)	5	2
Other service					13	N/A
No service code given					74	N/A

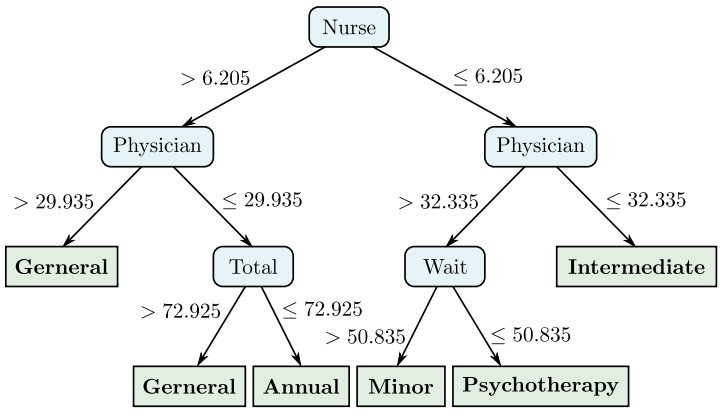


Figure 1.1: Consider the dataset studied by Bestvater et al. [20], where the authors focused on the impact of waiting times on patient's perception of service satisfaction. This survey collected the average time patients spend with the nurse and the physician for various services ranging from major and minor assessments to psychotherapy. We emphasize that the dataset was carefully curated to minimize leaking any *differentially private information* about the patients taking part in the survey. On the other hand, using a simple decision-tree classifier over this data, we found out that the timing data collected is leaking private information about patients in timing side-channels. For instance, if a patient spends less than 6 minutes with the nurse and spends close to 32 minutes with the physician with a low waiting time, the patient is visiting the hospital for a *psychotherapy* session!

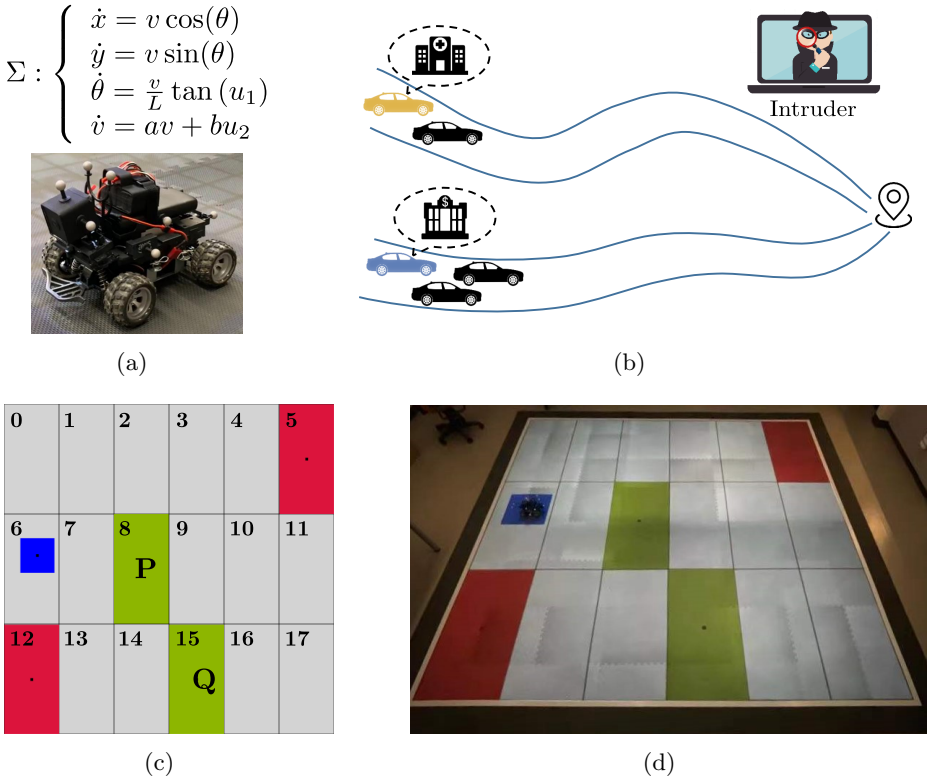


Figure 1.2: The AWS DeepRacer car and its dynamics (a); The plausible deniability of the car for secret initial region (b); The grid-world observations (c) where the red regions depict sensitive starting locations (e.g., hospital or bank) and the green regions represent the target; Our actual platform in the lab (d) corresponding to this grid-world.

developed as shown in Figure 1.2(d). Here we are interested in synthesizing a controller for the movement of a robotic vehicle (AWS DeepRacer Car in Figure 1.2(a)) with safety and security requirements. The intuition behind the security property of interest is as follows. Suppose the initial locations of the vehicle contain critical information which is needed to be kept secret, e.g., the vehicle might be a cash transit van that aims at transferring money initially from a bank to an ATM machine, or a patient who initially visited a hospital but unwilling to reveal this information to others. It is implicitly assumed that there is a malicious intruder who is observing the behavior of the vehicle remotely intending to carry out an attack. Therefore, it is in the interest of the system to verify whether it maintains plausible deniability for secret initial location where some confidential assignment is executed. In the physical platform, we assume that the vehicle can start from any of the four corner cell (Cells 0, 5, 12, and 17). We also assume that Cell 5 and Cell 12 marked in red are sensitive starting locations. We also assume that the time it takes for the robot to travel to any neighboring cell on east (E), west (W), north (N), and south (S) is the same and it is known to the intruder. Now assume

that the intruder can only observe when the robotic vehicle is in the regions marked by P (parking area) and Q (checkout queue) and gets the common observation G for the rest of the cells. A secure-by-construction controller synthesis task is to design a feedback controller satisfying the following requirements: 1) a mission requirement: the robotic vehicle visits regions P and Q infinitely often and 2) a privacy requirement: the intruder is unable to infer whether the vehicle got initiated from a sensitive location.

Suppose we design a controller providing control strategies from all initial cells such that the robot first follows a shortest path to reach Cell 8 or Cell 15, and then cycles between them forever. It is easy to verify that these control strategies satisfy the mission requirement of visiting regions P and Q infinitely often. However, unfortunately such controller does not satisfy the privacy requirement as it is clear from the following system executions adhering to the aforementioned control strategies: here on the left side we show the system executions, while on the right hand side we show the observations made by the intruder. The notation ω over parentheses shows the infinite repetition of the finite execution inside them.

- $0\xrightarrow{E}1\xrightarrow{E}2\xrightarrow{S}8(\xrightarrow{S}14\xrightarrow{E}15\xrightarrow{N}9\xrightarrow{W}8)^{\omega} \quad \mapsto G{\to}G{\to}G{\to}P(\to G{\to}Q{\to}G{\to}P)^{\omega}$
- $12\xrightarrow{E}13\xrightarrow{E}14\xrightarrow{N}8(\xrightarrow{S}14\xrightarrow{E}15\xrightarrow{N}9\xrightarrow{W}8)^{\omega} \quad \mapsto G{\to}G{\to}G{\to}P(\to G{\to}Q{\to}G{\to}P)^{\omega}$
- $5\xrightarrow{W}4\xrightarrow{W}3\xrightarrow{W}2\xrightarrow{S}8(\xrightarrow{S}14\xrightarrow{E}15\xrightarrow{N}9\xrightarrow{W}8)^{\omega} \quad \mapsto G{\to}G{\to}G{\to}G{\to}P(\to G{\to}Q{\to}G{\to}P)^{\omega}$
- $17\xrightarrow{W}16\xrightarrow{W}15(\xrightarrow{N}9\xrightarrow{W}8\xrightarrow{S}14\xrightarrow{E}15)^{\omega} \quad \mapsto G{\to}G{\to}Q(\to G{\to}P{\to}G{\to}Q)^{\omega}$

For this controller, if the system starts in the secret state 12, the corresponding observation is also matched by the non-secret state 0. On the other hand, when the system starts in secret state 5, there is no other non-secret initial state giving the same observation. Hence, whenever the system starts from the secret state 5, the observation uniquely identifies the initial state to be a secret one. For this controller, we say that the system is not opaque. On the other hand, by modifying the controller to change the strategy from Cell 17 to the one below makes the system opaque since it matches the observation sequence starting from Cell 5.

- $17\xrightarrow{N}11\xrightarrow{W}10\xrightarrow{W}9\xrightarrow{W}8(\xrightarrow{S}14\xrightarrow{E}15\xrightarrow{N}9\xrightarrow{W}8)^{\omega} \quad \mapsto G{\to}G{\to}G{\to}G{\to}P(\to G{\to}Q{\to}G{\to}P)^{\omega}$

A secure-by-construction synthesis framework aims to **automatically** design such controllers for large-scale CPS satisfying both the complex logic missions as well as the security requirements.

1.2 Contributions and Outline of the Thesis

This dissertation provides theoretical foundations to enable fast and reliable design of security-critical large-scale CPS by introducing a secure-by-construction, cost-efficient methodology. In **Chapter 2**, we present some mathematical notations and preliminaries from control theory which will be used throughout the thesis. Then, we introduce new notions of approximate opacity in **Chapter 3** for both stochastic and non-stochastic CPS. Our new definitions will enlarge the application domain of a security notion, called *opacity*, and serve as the foundations for the analysis or synthesis

of opacity for (stochastic) CPS. Based on the new notions of approximate opacity, we develop two different approaches for the analysis of security for complex CPS. One main approach for analyzing opacity of CPS is through abstraction-based techniques as discussed in **Chapter 4**. By introducing several new notions of approximate opacity-preserving simulation relations (or opacity-preserving stochastic simulation functions), we construct opacity-preserving finite abstractions which are finite systems that mimic the behaviors of concrete (stochastic) CPS in terms of opacity. This allows us to verify opacity of complex CPS using their finite abstractions. As an alternative to the abstraction-based framework, a deductive approach via a notion of barrier certificates is developed in **Chapter 5** for the analysis of approximate opacity for CPS. In order to reduce the computational complexity required in the proposed verification schemes, we also propose modular approaches as in **Chapter 6** to scale our secure-by-construction synthesis scheme for CPS by combining compositional synthesis techniques from computer science with those from control theory. **Chapter 7** concludes the results of the thesis and outlines potential future directions on related topics.

The main contribution of this dissertation is to provide mathematical foundations and efficient algorithms for the security analysis of large-scale CPS. To be specific, the contributions are as follows:

1) **Security notions for CPS (Chapter 3)**

   In order to provide a common framework for the analysis of complex CPS, we introduce various definitions from the discrete-event systems (DES), CPS, and formal methods communities. New notions of so-called approximate opacity are proposed for both stochastic and non-stochastic CPS. Approximate opacity captures the fact that CPS are usually metric systems in the sense that their outputs are physical signals rather than being discrete symbols. This new concept can be seen as a "robust" version of opacity by quantitatively characterizing the security guarantee level with respect to the measurement precision of the intruder. To provide a general setting for the secure-by-construction synthesis framework, we further unified the different requirements and security notions as a generalized language-based opacity notion.

   The covered materials in this chapter are based on the publications [114, 212, 115].

2) **Abstraction-based opacity verification of CPS (Chapter 4)**

   This chapter presents an abstraction-based opacity verification scheme for complex CPS. We provide for the first time a scheme for constructing opacity-preserving finite abstractions together with corresponding opacity-preserving simulation relations (or opacity-preserving stochastic simulation functions) for (stochastic) control systems. Moreover, we show that one can always construct opacity-preserving finite abstractions by leveraging the incremental input-to-state stability of (stochastic) control systems. Then, one can verify opacity of complex CPS by using their finite abstractions with a much cheaper computational cost. The results proposed here make the first step towards abstraction-based verification and synthesis of opacity. Furthermore, for the case of finite systems, we also

develop effective algorithms for the verification of new notions of approximate opacity, particularly, using *belief construction* techniques.

The materials presented in this chapter were published in [212, 115].

3) **A deductive approach for opacity verification via barrier certificates (Chapter 5)**

As an alternative to the abstraction-based approach developed in Chapter. 4, we present in this chapter a discretization-free deductive approach for the verification of opacity for CPS. We first define a notion of augmented systems that are constructed as the product of a control system and itself. Inspired by the duality of safety and reachability properties, we define a pair of so-called augmented control barrier certificates in conjunction with specified regions of interest capturing the initial and secret information of systems. The existence of the proposed barrier certificates for the augmented systems are shown to guarantee the (or the lack of) opacity for the original control systems. Although both barrier certificates only serve as sufficient conditions, they can be utilized in reverse directions in the sense that one ensures approximate opacity, and the other one shows the lack of approximate opacity of the control system. We also present a way to compute polynomial barrier certificates by means of sum-of-squares (SOS) programming under certain assumptions on the systems.

The results of this chapter appeared in the publications [117, 82].

4) **Modular verification of opacity for large-scale interconnected systems (Chapter 6)**

This chapter is devoted to mitigate the computational complexity issue tailored to the previously proposed verification approaches. Note that although the abstraction-based approach provided in Chapter 4 and the deductive approach presented in Chapter 5 are shown to be promising, a challenge lies in scaling the approaches for large-scale systems. In order to reduce the computational complexity, we propose a divide-and-conquer strategy to scale the proposed approaches by combining compositional synthesis techniques from computer science with those from control theory. Here, a large-scale system is tackled as an interconnection of smaller subsystems with manageable sizes. For the abstraction-based approach, instead of treating the interconnection monotonically, our compositionality result enables us to construct opacity-preserving finite abstractions for the subsystems individually. A top-down construction framework was presented equipped with a detailed algorithm as a guideline for the design of quantization parameters. Here, we propose compositionality results for both general interconnected control systems and interconnected switched systems which require different treatments on the state space discretization processes. For the barrier certificate approach, a compositional scheme for the construction of barrier certificates is also derived based on a small-gain type condition.

The covered materials in this chapter were published in [118, 113, 112, 81].

2 Preliminaries

In this chapter, we introduce some mathematical notations and preliminaries that will be used throughout this thesis. The presented preliminaries are based on the classical results from the fields of control theory, mathematics, and theoretical computer science.

2.1 Notation

We denote by $\mathbb{R}$, $\mathbb{Z}$, and $\mathbb{N}$ the set of real numbers, integer and non-negative integers, respectively. Additionally, these symbols are annotated with subscripts to restrict them in the usual way, e.g., the symbols $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$ and $\mathbb{N}_{>0}$ denote the set of non-negative real numbers, positive real numbers, and positive integers, respectively. We denote the closed, open, and half-open intervals in $\mathbb{R}$ by $[a\ b]$, $]a\ b[$, $[a\ b[$, and $]a\ b]$, respectively. For $a, b \in \mathbb{N}$ and $a \leq b$, we use $[a; b]$, $]a; b[$, $[a; b[$, and $]a; b]$ to denote the corresponding intervals in $\mathbb{N}$. Given $N \in \mathbb{N}_{\geq 1}$ vectors $\nu_i \in \mathbb{R}^{n_i}$, $n_i \in \mathbb{N}_{\geq 1}$, and $i \in [1; N]$, we write $\nu = (\nu_1, \ldots, \nu_N)$ to denote the corresponding concatenated vector in $\mathbb{R}^n$ with $n = \sum_i n_i$. Given a vector $x \in \mathbb{R}^n$, we denote the infinity norm of x by $\|x\|$ and the Euclidean norm of x by $\|x\|_2$. Given any $a \in \mathbb{R}$, $|a|$ denotes the absolute value of a. We denote by id the identity function over $\mathbb{R}$, i.e., $\mathrm{id}(r) = r$ for all $r \in \mathbb{R}$.

Given two sets $X, Y \subseteq \mathbb{R}^n$, the complement of set X with respect to Y is defined as $Y \backslash X = \{x : x \in Y, x \notin X\}$. For any set $Z \subseteq \mathbb{R}^n$, ∂Z and $\overline{Z}$, respectively, denotes the boundary and topological closure of Z. The Minkowski sum of two sets $X, Y \subseteq \mathbb{R}^n$ is defined by $X \oplus Y = \{z \in \mathbb{R}^n : \exists_{x \in X, y \in Y}, z = x + y\}$. Given a function $f : \mathbb{N}_{\geq 0} \to \mathbb{R}^n$, the (essential) supremum of f is denoted by $\|f\|_\infty := (\mathrm{ess})\sup\{\|f(k)\|, k \geq 0\}$. We identify a relation $R \subseteq A \times B$ with the map $R : A \to 2^B$ defined by $b \in R(a)$ iff $(a, b) \in R$. Given a relation $R \subseteq A \times B$, R^{-1} denotes the inverse relation defined by $R^{-1} = \{(b, a) \in B \times A : (a, b) \in R\}$. The closed ball centered at $u \in \mathbb{R}^m$ with radius λ is defined by $\mathcal{B}_\lambda(u) = \{v \in \mathbb{R}^m : \|u - v\| \leq \lambda\}$. We denote the closed ball centered at the origin in $\mathbb{R}^n$ and with radius λ by $\mathcal{B}_\lambda$. A set $B \subseteq \mathbb{R}^m$ is called a *box* if $B = \prod_{i=1}^m [c_i, d_i]$, where $c_i, d_i \in \mathbb{R}$ with $c_i < d_i$ for each $i \in \{1, \ldots, m\}$. Geometrically, for any $\mu \in \mathbb{R}^+$ with $\mu \leq \mathit{span}(B)$ and $\lambda \geq \mu$, the collection of sets $\{\mathcal{B}_\lambda(p)\}_{p \in [B]_\mu}$ is a finite covering of B, i.e., $B \subseteq \bigcup_{p \in [B]_\mu} \mathcal{B}_\lambda(p)$. For any set $S \subseteq \mathbb{R}^n$ of the form of finite union of boxes, e.g., $S = \bigcup_{j=1}^M S_j$ for some $M \in \mathbb{N}$, where $S_j = \prod_{i=1}^n [c_i^j, d_i^j] \subseteq \mathbb{R}^n$ with $c_i^j < d_i^j$, we define $\mathit{span}(S) = \min_{j=1,\ldots,M} \eta_{S_j}$ and $\eta_{S_j} = \min\{|d_1^j - c_1^j|, \ldots, |d_n^j - c_n^j|\}$. Moreover, for a set in the form of $X = \prod_{i=1}^N X_i$, where $X_i \subseteq \mathbb{R}^{n_i}$ are of the form of finite union of boxes, and any positive (component-wise) vector $\eta = (\eta_1, \ldots, \eta_N)$ with $\eta_i \leq \mathit{span}(X_i)$, $\forall i \in [1; N]$, we define $[X]_\eta = \prod_{i=1}^N [X_i]_{\eta_i}$, where $[X_i]_{\eta_i} = [\mathbb{R}^{n_i}]_{\eta_i} \cap X_i$ and $[\mathbb{R}^{n_i}]_{\eta_i} = \{a \in \mathbb{R}^{n_i} : a_j = k_j \eta_i, k_j \in \mathbb{Z}, j = 1, \ldots, n_i\}$. Remark that $[X]_\eta \neq \varnothing$ for

any $\eta \leq span(X)$. Given a set $\mathbb{S} \subseteq \mathbb{R}^n$ and a constant $\theta \in \mathbb{R}_{\geq 0}$, we define a new set $\mathbb{S}^\theta = \mathbb{S} \oplus \mathcal{B}_\theta$ as the inflated version of set $\mathbb{S}$.

We use notations $\mathcal{K}$, $\mathcal{K}_\infty$, and $\mathcal{KL}$ to denote the different classes of comparison functions, as follows: $\mathcal{K} = \{\gamma : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} : \gamma \text{ is continuous, strictly increasing and } \gamma(0) = 0\}$; $\mathcal{K}_\infty = \{\gamma \in \mathcal{K} : \lim_{r\to\infty} \gamma(r) = \infty\}$; $\mathcal{KL} = \{\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} : \text{for each fixed } s, \text{ the map } \beta(r, s) \text{ belongs to class } \mathcal{K} \text{ with respect to } r \text{ and, for each fixed nonzero } r, \text{ the map } \beta(r, s) \text{ is decreasing with respect to } s \text{ and } \beta(r, s) \to 0 \text{ as } s \to \infty\}$.

For a set A, we write A^* for the set of finite sequences with elements in A and A^ω for the set of (infinite) ω-sequences. We write $A^\infty = A^* \cup A^\omega$.

2.2 System Model

Transition systems are natural, expressive, unified, and widely accepted [186, 17, 13] semantics for cyber-physical systems describing both continuous-space and finite control systems.

Definition 2.2.1. (**Transition Systems**) A transition system Σ is described by a quadruple

$$\Sigma = (X, X_0, U, \longrightarrow), \tag{2.2.1}$$

where X is a (possibly infinite) set of states, $X_0 \subseteq X$ is a (possibly infinite) set of initial states, U is a (possibly infinite) set of inputs, and $\longrightarrow \subseteq X \times U \times X$ is a transition relation. We call a system *finite* (or *symbolic*), if X and U are finite sets.

A transition $(x, u, x') \in \longrightarrow$ is also denoted by $x \xrightarrow{u} x'$. For a transition $x \xrightarrow{u} x'$, state x' is called a u-successor, or simply a successor, of state x; state x is called a u-predecessor, or simply a predecessor, of state x'. We denote by $\mathbf{Post}_u(x)$ the set of all u-successors of state x and by $\mathbf{Pre}_u(x)$ the set of all u-predecessors of state x. For a set of states $q \in 2^X$, we write

$$\mathbf{Post}_u(q) = \cup_{x\in q}\mathbf{Post}_u(x) \text{ and } \mathbf{Pre}_u(q) = \cup_{x\in q}\mathbf{Pre}_u(x).$$

We call a system *deterministic*, if for any state $x \in X$ and any input $u \in U$, $\mathbf{Post}_u(x)$ is a singleton or an empty set; otherwise we call it *non-deterministic*.

A system Σ starting from an initial state $x_0 \in X_0$ and under input sequence $u_1u_2\cdots u_n \in U^*$, induces a finite state *run*

$$x_0 \xrightarrow{u_1} x_1 \xrightarrow{u_2} \cdots \xrightarrow{u_{n-1}} x_{n-1} \xrightarrow{u_n} x_n, \tag{2.2.2}$$

such that $x_i \xrightarrow{u_{i+1}} x_{i+1}$ for all $0 \leq i < n$. Note that the run induced by an input sequence may not be unique because the system may be non-deterministic.

We call a finite sequence of states $x_0x_1\cdots x_n \in X^*$ a *finite path* of the system Σ and denote by $\texttt{Path}(\Sigma, x_0)$ the set of all finite paths generated by Σ starting from x_0 and with $\texttt{Path}(\Sigma) = \cup_{x_0\in X_0}\texttt{Path}(\Sigma, x_0)$. Similarly, an infinite path $x_0x_1\cdots \in X^\omega$ is

an ω-sequence defined analogously and we denote by $\texttt{Path}^\omega(\Sigma, x_0)$ the set of all infinite paths of Σ from x_0 and with $\texttt{Path}^\omega(\Sigma) = \cup_{x_0 \in X_0} \texttt{Path}^\omega(\Sigma, x_0)$.

Behaviors A primary concern is whether the behaviors of system Σ satisfy some desired specification. Formally, let $\mathcal{AP}$ be a finite set of features, or *(atomic) propositions*, of the state space. We view the states with the lenses of atomic propositions, and to do so, we define a *labeling function* $L : X \to 2^{\mathcal{AP}}$ that assigns to each state $x \in X$ in Σ a set of propositions $L(x)$ true at the state x. The labeling function can naturally be extended from states to path: we call such labeling of a path a *trace*. For any finite or infinite path $\mathbf{x} = x_0 x_1 \cdots \in X^\infty$, its trace is $L(\mathbf{x}) = L(x_0)L(x_1)\cdots \in (2^{\mathcal{AP}})^\infty$. The set of all finite traces and the set of all infinite traces are denoted by $\texttt{Trace}(\Sigma)$ and $\texttt{Trace}^\omega(\Sigma)$, respectively.

Observations The system releases information to the external world during its execution. The external world often may not observe the internal states X or their atomic propositions directly but rather their properties over some observation symbols. Let Y be such set of observations. Let the *output function* $h : X \to Y$ determine the external observation of each internal state $x \in X$. It can naturally be extended to finite or infinite paths, i.e., for a path $\mathbf{x} = x_0 x_1 \cdots \in X^\infty$, its *output* corresponds to a sequence $h(\mathbf{x}) = h(x_0)h(x_1)\cdots \in Y^\infty$.

The system Σ is said to be *metric* if the observation set Y is equipped with a metric $\mathbf{d} : Y \times Y \to \mathbb{R}_{\geq 0}$. For any two paths $\mathbf{x} = x_0 x_1 \cdots$ and $\mathbf{x}' = x_0' x_1' \cdots$, we say the outputs of $\mathbf{x}$ and $\mathbf{x}'$ are *(exactly) output equivalent*, denoted by $h(\mathbf{x}) = h(\mathbf{x}')$, if $h(x_i) = h(x_i')$ for all $i \geq 0$; on the other hand, we say that they are *δ-approximately output equivalent*, and write $h(\mathbf{x}) \approx_\delta h(\mathbf{x}')$, if $\sup_{i \geq 0} \mathbf{d}(h(x_i), h(x_i')) \leq \delta$.

To emphasize the labeling $L : X \to 2^{\mathcal{AP}}$ and output functions $h : X \to Y$ of a system Σ, we rewrite the tuple describing the system as

$$\Sigma = (X, X_0, U, \longrightarrow, \mathcal{AP}, L, Y, h).$$

When it is clear from the context, we may drop some of the elements in the tuple for the sake of simple presentation.

Remark 2.2.2. *In the DES literature, it is customary to model a system as a finite state machine $G = (X, E, \delta, X_0, E_o)$, where X is a set of states, E is a set of events, $\delta : X \times E \to 2^X$ is a transition function and $X_0 \subseteq X$ is a set of initial states [26]. In such treatments, both inputs and properties are captured by events E. Furthermore, it is also assumed that the observation mapping is also event-based captured by a natural projection $P : E \to E_o$.*

Our modeling framework is general enough to capture treatment in DES literature and capable of expressing more general scenarios posed in the reactive control systems settings.

2.3 Discrete-Time Control Systems

In this thesis, we consider *control* systems in discrete time, which is a metric system as discussed in Section 2.2. In the remainder of this thesis, we assume that the output set Y is equipped with the infinity norm: $\mathbf{d}(y_1, y_2) = \|y_1 - y_2\|$, $\forall y_1, y_2 \in Y$. We have a similar assumption for the state set X. Note that in the remainder of this dissertation, we will mainly consider control systems with secret states (cf. Section 3.2) and incorporate the secrete state set $X_S \subseteq X$ in the system definitions. The formal definition of discrete-time control systems is given as follows.

Definition 2.3.1. (**Discrete-Time Control Systems**) A discrete-time control system (dt-CS) Σ is defined by the tuple

$$\Sigma = (X, X_0, X_S, U, f, Y, h), \tag{2.3.1}$$

where $X \subseteq \mathbb{R}^n$, $U \subseteq \mathbb{R}^m$, and $Y \subseteq \mathbb{R}^q$ are the state, input, and output sets, respectively. Sets $X_0, X_S \subseteq X$ are the sets of initial states and secret states, respectively. The map $f : X \times U \to X$ is called the transition function, and $h : X \to Y$ is the output map and assumed to satisfy the following Lipschitz condition: $\|h(x) - h(y)\| \leq \alpha(\|x - y\|)$ for some $\alpha \in \mathcal{K}_\infty$ and all $x, y \in X$. The dynamics of Σ is described by difference inclusions of the form

$$\Sigma : \begin{cases} \mathbf{x}(k+1) \in & f(\mathbf{x}(k), \nu(k)), \\ \mathbf{y}(k) = & h(\mathbf{x}(k)), \end{cases} \tag{2.3.2}$$

where $k \in \mathbb{N}$, $\mathbf{x} : \mathbb{N} \to X$, $\mathbf{y} : \mathbb{N} \to Y$, and $\nu : \mathbb{N} \to U$ are the state, output, and input signals, respectively.

We write $\mathbf{x}_{x_0,\nu}(k)$ to denote the point reached at time k under the input signal ν from initial condition x_0. Similarly, we denote by $\mathbf{y}_{x_0,\nu}(k)$ the output corresponding to state $\mathbf{x}_{x_0,\nu}(k)$, i.e., $\mathbf{y}_{x_0,\nu}(k) = h(\mathbf{x}_{x_0,\nu}(k))$. In the above definition, we implicitly assumed that set X is positively invariant[1].

2.3.1 Discrete-Time Stochastic Control Systems

In some part of this thesis, we consider discrete-time stochastic control systems formally defined as follows.

Definition 2.3.2. (**Discrete-Time Stochastic Control Systems**) A discrete-time stochastic control systems (dt-SCS) is defined by the tuple

$$\Sigma = (X, X_0, X_S, U, \varsigma, f, Y, h)\,, \tag{2.3.3}$$

where $X \subseteq \mathbb{R}^n$, $U \subseteq \mathbb{R}^m$, and $Y \subseteq \mathbb{R}^q$ are Borel sets denoting the state, input and output sets of the system, respectively. Sets $X_0, X_S \subseteq X$ are the sets of initial and

[1] Set X is called positively invariant under (2.3.2) if $\mathbf{x}_{x_0,\nu}(k) \in X$ for any $k \in \mathbb{N}$, any $x_0 \in X$ and any $\nu : \mathbb{N}_0 \to U$.

secret states, respectively. We use $\mathcal{B}(X)$ to denote the Borel sigma-algebra on the state set X, thus $(X, \mathcal{B}(X))$ denotes the corresponding measurable space. In the probability space $(\Omega, \mathcal{F}_\Omega, \mathbb{P})$, we use $\varsigma = (\varsigma(1), \varsigma(2), \dots)$ to denote a sequence of independent and identically distributed (i.i.d.) random variables from Ω to the measurable set V_ς, where $\varsigma(k) : \Omega \to V_\varsigma, k \in \mathbb{N}$. The maps $f : X \times U \times V_\varsigma \to X$ and $h : X \to Y$ are measurable functions serving as the state transition relation and output map, respectively. Given an initial state $\xi(0) \in X$ and $\forall k \in \mathbb{N}$, the dt-SCS Σ satisfies

$$\Sigma : \begin{cases} \xi(k+1) = f(\xi(k), \nu(k), \varsigma(k)), \\ \zeta(k) = h(\xi(k)), \end{cases} \tag{2.3.4}$$

where $\xi(\cdot) : \mathbb{N} \to X$, $\zeta(\cdot) : \mathbb{N} \to Y$, and $\nu(\cdot) : \mathbb{N} \to U$ are the state, output, and input signals, respectively. We use $\mathcal{U}$ to denote a collection of sequences $\nu : \Omega \to U$, where $\nu(k)$ is independent of $\varsigma(t)$ for any $k, t \in \mathbb{N}$ and $t \geq k$.

A dt-SCS defined in (2.3.3) with (possibly) continuous state set can be equivalently represented as a general Markov decision process (gMDP). We refer the interested readers to [59] for formal definitions of gMDPs. Note that by capturing stochastic systems as gMDP, this modeling framework is general enough to include as special cases labelled Markov processes [42], discrete-time stochastic hybrid games [43], stochastic switched systems [102], and so on.

Given system $\Sigma = (X, X_0, X_S, U, \varsigma, f, Y, h)$, $x \xrightarrow{u} x'$ is called a *transition* in the system if and only if $x' = f(x, u, \varsigma)$. The random sequence $\xi_{x_0\nu} : \Omega \times \mathbb{N} \to X$, which is in the form of $\xi_{x_0\nu} = (x_0, x_1, \dots, x_n)$, is said to be a *solution process* of Σ under input sequence $\nu = (u_1, u_2, \dots, u_n)$ satisfying (2.3.4), with initial state $\xi_{x_0\nu}(0) = x_0$. The random sequence $\zeta_{x_0\nu} : \Omega \times \mathbb{N} \to Y$ is called the *output run* and defined as $\zeta_{x_0\nu} = (y_0, y_1, \dots, y_n)$ such that there exists a solution process $\xi_{x_0\nu} = (x_0, x_1, \dots, x_n)$ with $y_i = h(x_i)$, for $i \in \{0, \dots, n\}$. A solution process and a finite output run can be extended to an infinite state run and an infinite output run as well.

2.3.2 Discrete-Time Switched Systems

In some part of the thesis, we consider discrete-time switched systems of the following form.

Definition 2.3.3. (**Discrete-Time Switched Systems**) A discrete-time switched system (dt-SS) Σ is defined by the tuple $\Sigma = (\mathbb{X}, \mathbb{X}_0, \mathbb{X}_S, P, F, Y, h)$, where $\mathbb{X} \subseteq \mathbb{R}^n, Y \subseteq \mathbb{R}^q$ are the state and output sets, respectively. Sets $\mathbb{X}_0, \mathbb{X}_S \subseteq \mathbb{X}$ are the sets of initial states and secret states, respectively. Set $P = \{1, \dots, m\}$ is a finite set of modes, $F = \{f_1, \dots, f_m\}$ is a collection of set-valued maps $f_p : X \rightrightarrows X$ for all $p \in P$, and $h : X \to Y$ is the output map. The dt-SS Σ is described by difference inclusions of the form

$$\Sigma : \begin{cases} \mathbf{x}(k+1) & \in f_{\mathsf{p}(k)}(\mathbf{x}(k)), \\ \mathbf{y}(k) & = h(\mathbf{x}(k)), \end{cases} \tag{2.3.5}$$

where $k \in \mathbb{N}$, $\mathbf{x} : \mathbb{N} \to \mathbb{X}$, $\mathbf{y} : \mathbb{N} \to Y$, and $\mathsf{p} : \mathbb{N} \to P$ are the state, output, and switching signal, respectively.

Let $\varphi_k, k \in \mathbb{N}_{\geq 1}$, denote the time when the k-th switching instant occurs. We assume that signal p satisfies a dwell-time condition [106] (i.e. there exists $k_d \in \mathbb{N}_{\geq 1}$, called the dwell-time, such that for all consecutive switching time instants φ_k, φ_{k+1}, $\varphi_{k+1} - \varphi_k \geq k_d$). We assume that for every initial condition and any sequence of switching signals, the corresponding state signal is defined for all $k \geq 0$.

Moreover, we employ the notion of *transition systems* as discussed in Definition 2.2.1, to provide an alternative description of switched systems that can be later directly related to their finite abstractions in a common framework.

Definition 2.3.4. *(Discrete-Time Switched Systems as Transition Systems) Given a dt-SS $\Sigma = (\mathbb{X}, \mathbb{X}_0, \mathbb{X}_S, P, F, Y, h)$, we define the associated transition system $T(\Sigma) = (X, X_0, X_S, U, \mathcal{F}, Y, \mathcal{H})$, where:*

- $X = \mathbb{X} \times P \times \{0, \ldots, k_d - 1\}$ *is the state set;*
- $X_0 = \mathbb{X}_0 \times P \times \{0\}$ *is the initial state set;*
- $X_S = \mathbb{X}_S \times P \times \{0, \ldots, k_d - 1\}$ *is the secret state set;*
- $U = P$ *is the input set;*
- $\mathcal{F}$ *is the transition function given by $(x^+, p^+, l^+) \in \mathcal{F}((x, p, l), u)$ if and only if $x^+ \in f_p(x), u = p$ and one of the following scenarios hold:*
  - *$l < k_d - 1$, $p^+ = p$ and $l^+ = l + 1$: switching is not allowed because the time elapsed since the latest switch is strictly smaller than the dwell time;*
  - *$l = k_d - 1$, $p^+ = p$ and $l^+ = k_d - 1$: switching is allowed but no switch occurs;*
  - *$l = k_d - 1$, $p^+ \neq p$ and $l^+ = 0$: switching is allowed and a switch occurs;*
- $Y = Y$ *is the output set;*
- $\mathcal{H} : X \to Y$ *is the output map defined as $\mathcal{H}(x, p, l) = h(x)$.*

Note that in the above definition, two additional variables p and l are added to the state tuple of the system Σ. The variable l serves as a counter to record the sojourn time of the switching signal, which allows or prevents the system from switching depending on whether the dwell-time condition is satisfied; the variable p acts as a memory to record the current mode of the system.

The following proposition is borrowed from [184] showing that the output runs of a dt-SS Σ and its associated transition system $T(\Sigma)$ are equivalent so that one can use Σ and $T(\Sigma)$ interchangeably.

Proposition 2.3.5. *Consider a transition system $T(\Sigma)$ in Definition 2.3.4 associated to Σ as in Definition 2.3.3. Any output trajectory of Σ can be uniquely equated matched with an output trajectory of $T(\Sigma)$ and vice versa.*

2.4 Incremental Input-to-State Stability

In this section, we introduce preliminary results on the notion of incremental input-to-state stability (δ-ISS) which will be leveraged later to show some of the main results of this thesis.

Definition 2.4.1. *[193] A discrete-time control system $\Sigma = (X, X_0, X_S, U, f, Y, h)$ is called incrementally input-to-state stable (δ-ISS) if there exist a $\mathcal{KL}$ function β and $\mathcal{K}_\infty$ function γ such that $\forall x, x' \in X$ and $\forall \nu, \nu' : \mathbb{N}_0 \to U$, the following inequality holds for any $k \in \mathbb{N}$:*

$$\|\mathbf{x}_{x,\nu}(k) - \mathbf{x}_{x',\nu'}(k)\| \leq \beta(\|x - x'\|, k) + \gamma(\|\nu - \nu'\|_\infty). \tag{2.4.1}$$

Example 2.4.2. *As an example, a linear control system:*

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + B\nu(k), \quad \mathbf{y}(k) = C\mathbf{x}(k), \tag{2.4.2}$$

is δ-ISS if all eigenvalues of A are inside the unit circle. In this case, functions β and γ can be chosen as:

$$\beta(r, k) = \|A^k\| r; \quad \gamma(r) = \|B\| \left(\sum_{m=0}^{\infty} \|A^m\| \right) r. \tag{2.4.3}$$

In general, it is difficult to check inequality (2.4.1) directly for nonlinear systems. Fortunately, δ-ISS can be characterized using Lyapunov functions.

Definition 2.4.3. *[193] Consider a discrete-time control system Σ and a continuous function $V : X \times X \to \mathbb{R}_{\geq 0}$. Function V is called a δ-ISS Lyapunov function for Σ if there exist $\mathcal{K}_\infty$ functions $\underline{\alpha}, \overline{\alpha}, \rho$ and $\mathcal{K}$ function σ such that:*

(i) for any $x, x' \in X$
$\underline{\alpha}(\|x - x'\|) \leq V(x, x') \leq \overline{\alpha}(\|x - x'\|)$;

(ii) for any $x, x' \in X$ and $u, u' \in U$
$V(f(x, u), f(x', u')) - V(x, x') \leq -\rho(V(x, x')) + \sigma(\|u - u'\|)$;

The following result characterizes δ-ISS in terms of existence of δ-ISS Lyapunov functions.

Theorem 2.4.4. *[193] Consider a control system Σ. System Σ is δ-ISS if it admits a δ-ISS Lyapunov function.*

The next technical lemma will be used later to show some of the main results of next chapters.

Lemma 2.4.5. *Consider a control system Σ. Suppose V is a δ-ISS Lyapunov function for Σ. Then there exist $\kappa, \lambda \in \mathcal{K}_\infty$, where $\kappa(s) < s$ for any $s \in \mathbb{R}^+$, such that*

$$V(f(x, u), f(x', u')) \leq \max\{\kappa(V(x, x')), \lambda(\|u - u'\|)\}, \tag{2.4.4}$$

for any $x, x' \in X$ and any $u, u' \in U$.

The proof is similar to that of Theorem 1 in [182] and is omitted here.

3 Security Notions for Cyber-physical Systems

3.1 Introduction

Security requirements, in the DES [107, 202, 209, 95] and control theory communities, are often expressed using the notion of *opacity*, while in the realm of computer science security requirements are expressed using closely related, but subtly different, concepts of non-interference [128, 134, 201], K-safety [178, 139], language-based secrecy [3], and their generalizations using HyperLTL properties [33, 32].

In this chapter, we review the above-mentioned notions and provide a generalized security notion for CPS. In particular, we revisit various definitions and synthesize ideas from three research communities: discrete event systems, control systems, and formal methods to pose and study central problems supporting secure-by-construction synthesis. In our selection, the focus of the DES community is on the finite state models, the control systems community primarily on the continuous space models, while the results from formal methods community will primarily focus on logical and automata-theoretic results. Then, in Section 3.3, new notions of opacity are developed to suitably capture the metric nature of output sets of real-world (stochastic) CPS. We then provide a unifying view of various models and problems studied in this context, which integrate works in the three research fields under a common general framework.

3.2 Security Notions for Finite Systems: Opacity

Attack Model. In the setting discussed here, we assume that there exists a *secret predicate* on runs that models the confidential behavior of the system. The system does not want the intruder to infer the status of the secret predicate, i.e., whether it has executed a secret run. We consider that the intruder knows the dynamics of the system; and can observe the output sequences of the system. The intruder wants to use the output sequences observed online and the knowledge of the system model to infer certain information about the secret predicates of the corresponding run. For simplicity, we assume that the input sequences are internal information and unknown to the intruder. This setting can be easily relaxed to handle the case where both input and output information are available to the intruder.

Opacity is a well-studied confidentiality property that captures whether or not the "secret" of the system can be revealed to an intruder that can infer the system's actual behavior based on the information flow. A system is said to be opaque if it always has the plausible deniability for any of its secret behavior.

Definition 3.2.1 (**Language-Based Opacity**)**.** For a system $\Sigma = (X, X_0, U, \longrightarrow, Y, h)$, let $\mathcal{P}_S \subseteq \texttt{Path}(\Sigma)$ be the set of secret (finite) paths and $\mathcal{P}_P \subseteq \texttt{Path}(\Sigma)$ be a set of non-secret (finite) paths. We say system Σ is **opaque** w.r.t. $\mathcal{P}_S$ and $\mathcal{P}_P$ if for any secret path $\mathbf{x} \in \mathcal{P}_S$, there exists a non-secret path $\mathbf{x}' \in \mathcal{P}_P$ such that $h(\mathbf{x}) = h(\mathbf{x}')$.

The above definition of opacity is referred to as *language-based opacity* in the DES literature [107] as it uses languages $\mathcal{P}_S$ and $\mathcal{P}_P$ to represent secret and non-secret behaviors, respectively. The condition in the definition can also be equivalently written in terms of language inclusion as follows:

$$h(\mathcal{P}_S) \subseteq h(\mathcal{P}_P). \tag{3.2.1}$$

In specific applications, secret paths $\mathcal{P}_S$ usually have concrete meanings, e.g., currently at a secret location or initiated from a secret location. Therefore, a commonly used approach is to consider a set of *secret states* $X_S \subseteq X$. Depending on what information the system wants to hide, the following *state-based* notions of opacity have been introduced in the literature. In the remainder part of this dissertation, when discussing state-based opacity, we incorporate the secrete state set X_S in the system definition.

Definition 3.2.2 (**State-Based Opacity**)**.** Let $\Sigma = (X, X_0, X_S, U, \longrightarrow, Y, h)$ be a system, and $K \in \mathbb{N}$ be a non-negative integer. We say system Σ is

- *Initial-State Opaque* [162] if for any path $\mathbf{x} = x_0 x_1 \cdots x_n \in \texttt{Path}(\Sigma)$, where $x_0 \in X_S$, there exists a path $\mathbf{x}' = x'_0 x'_1 \cdots x'_n \in \texttt{Path}(\Sigma)$, where $x'_0 \notin X_S$, such that $h(\mathbf{x}) = h(\mathbf{x}')$;
- *Current-State Opaque* [160] if for any path $\mathbf{x} = x_0 x_1 \cdots x_n \in \texttt{Path}(\Sigma)$, where $x_n \in X_S$, there exists a path $\mathbf{x}' = x'_0 x'_1 \cdots x'_n \in \texttt{Path}(\Sigma)$, where $x'_n \notin X_S$, such that $h(\mathbf{x}) = h(\mathbf{x}')$;
- *Infinite-Step Opaque* [165] if for any path $\mathbf{x} = x_0 x_1 \cdots x_n \ldots x_{n+k} \in \texttt{Path}(\Sigma)$, where $x_n \in X_S$, there is a path $\mathbf{x}' = x'_0 x'_1 \cdots x'_n \ldots x'_{n+k} \in \texttt{Path}(\Sigma)$, where $x'_n \notin X_S$, such that $h(\mathbf{x}) = h(\mathbf{x}')$;
- *K-Step Opaque* [161] if for any path $\mathbf{x} = x_0 x_1 \cdots x_n \ldots x_{n+k} \in \texttt{Path}(\Sigma)$, where $x_n \in X_S$ and $k \leq K$, there exists a path $\mathbf{x}' = x'_0 x'_1 \cdots x'_n \cdots x'_{n+k} \in \texttt{Path}(\Sigma)$, where $x'_n \notin X_S$, such that $h(\mathbf{x}) = h(\mathbf{x}')$;
- *Pre-Opaque* [207] if for any path $\mathbf{x} = x_0 x_1 \cdots x_n$ and any $k \in \mathbb{N}$, there exists a path $\mathbf{x}' = x'_0 x'_1 \cdots x'_n \cdots x'_{n+k} \in \texttt{Path}(\Sigma)$, where $x'_{n+k} \notin X_S$, such that $h(x_0 x_1 \ldots x_n) = h(x'_0 x'_1 \cdots x'_n)$.

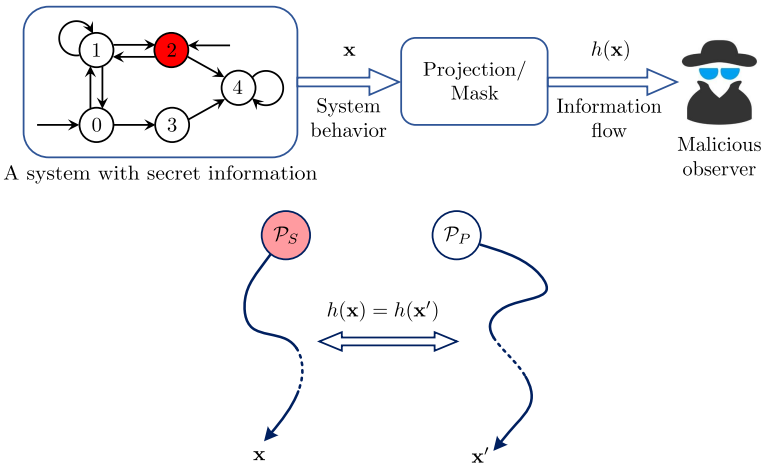


Figure 3.1: Graphical illustration of initial-state opacity.

The above state-based notions of opacity are closely related to the three fundamental state estimation problems in the systems theory: filtering, smoothing and prediction [57]. Specifically, current-state opacity is related to the *filtering* problem because it requires that the intruder can never determine for sure that the system is currently at a secret state. Initial-state opacity and infinite/K-step opacity are related to the *smoothing* problem because they both consider the scenario where the intruder can use latter observations to infer whether or not a system was at a secret state for some previous or the initial instant. In particular, initial-state opacity says that the intruder can never know that the system was initiated from a secret state, and K-step opacity says that the intruder can never know that the system was at a secret state within the past K-steps. Clearly, when K takes values 0 and ∞, K-step opacity becomes current-state opacity and infinite-step opacity, respectively. Finally, the notion of pre-opacity is related to the *predication* problem by requiring that the intruder can never know for sure that the system will reach a secret state for some specific future instant. This type of opacity essentially captures the intention security of the system. As an example, an illustration of the concept of initial-state opacity is depicted in Figure 3.1.

Note that our definition of infinite-step opacity requires that the intruder should never know for sure that the system is/was at a secret state *for any specific instant*. In some cases, the intruder may know that the system must have visited a secret state, although it cannot tell the precise instant. This requirement can be captured by the notion of strong (or trajectory-based) infinite-step opacity; see, e.g., Remark 5 in [165]. This notion is stronger than ours and which one to use is dependent on the applications. However, strong infinite-step opacity can be transformed to current-state opacity by augmenting the state-space to encode whether or not a secret state has been visited.

More recently, the definitions and verification algorithms for different notions of opacity have been extended to other classes of (discrete) systems, including Petri nets [191, 192, 35, 16], stochastic systems [163, 84, 200], recursive tile systems [29] and

pushdown systems [91]. The interested readers are referred to recent surveys [75, 95] for more references and recent developments on this active research area.

3.3 Security Notions for CPS: Approximate Opacity

Since opacity is an information-flow property, its definition strictly depends on the information model of the system. The formulation of opacity in the last subsection requires that for any secret behavior, there exists a non-secret behavior such that they generate exactly the same output. Therefore, we will also refer to these definitions as *exact opacity*. Exact opacity essentially assumes that the intruder or the observer can always measure each output or distinguish between two different outputs precisely. This setting is reasonable for non-metric systems where outputs are symbols or events. However, for metric systems, e.g., when the outputs are physical signals, this setting is too restrictive. In particular, due to the imperfect measurement precision, which is almost the case for all physical systems, it is very difficult to distinguish two observations if their difference is very small. A typical example of this scenario is linear or nonlinear discrete-time control systems with continuous state-spaces and continuous output mappings. Therefore, exact opacity may be too strong for metric systems and it is meaningful to define a weak and "robust" version of opacity.

In this section, we develop novel concepts of opacity called *approximate opacity* that are suitable for (stochastic) CPS possibly with continuous sets of states, inputs, and outputs. These concepts are proposed to quantitatively evaluate the security guarantee level with respect to the measurement precision of the intruder. The new concepts can be seen as a "robust" version of opacity by characterizing under what measurement precision the system is opaque.

Related Works The problem studied in this section is closely related to several works in the literature. First, several different approaches have been proposed in the literature to evaluate opacity more quantitatively rather than requiring that the system is opaque exactly [163, 19, 30, 211]. For example, in [30], the authors adopt the Jensen-Shannon divergence as the measurement to quantify secrecy loss. For finite systems, one popular approach is to consider systems modeled by probabilistic finite-state automata, Markov chains or Markov decision processes. Then one can quantify opacity in terms of probability [163, 18, 19, 84, 30, 211, 103]. These approaches essentially aim to analyze how opaque a single system is, e.g., the probability of being opaque. However, they neither consider how close two systems are in terms of being opaque nor consider under what observation precision level, we can guarantee opacity.

There are also attempts in the literature that extend opacity from discrete systems to continuous systems. For example, in the recent results in [154, 153, 155], the authors extended the notion of opacity to (switched) linear systems. However, their definition of opacity is more related to an output reachability property rather than an information-flow property. Moreover, their formulation is mostly based on the setting of exact opacity, i.e., we can always distinguish two different outputs precisely no matter how

close they are. In [154], the authors mentioned the direction of using output metric to quantify opacity and a property called strong ϵ-$\mathcal{K}$-initial-state opacity was proposed, which is closely related to our notions. However, no systematic study, e.g., verification and abstraction as we consider in this thesis, was provided for this property. A related notion called *differential privacy* was introduced in [44] for database systems and has attracted significant attention in the past few years [99, 79, 208]. In particular, [79] extends the original notion of differential privacy to symbolic systems. Differential privacy requires that any two adjacent data should produce indistinguishable outputs in the probability sense. However, the essence of opacity is a confidentiality property that captures the plausible deniability of the system's secret behavior, while differential privacy captures whether or not any sensitive data can be learned under some side-information. These two properties are incomparable in general. Finally, approximate notions of two related properties called *diagnosability* and *predictability* are investigated recently in [142, 47]. Their setting is very similar to us as we both consider a measurement uncertainty threshold. However, diagnosability and predictability can be preserved by standard approximate simulation relation, wwhereas it was proved in [222] that standard approximate simulation relation does not preserve opacity. These notions are again different in essence.

3.3.1 Approximate Opacity for Non-Stochastic Control Systems

In this subsection, we define a concept called *approximate opacity* for non-stachastic control systems. In particular, we treat two outputs as "indistinguishable" outputs if their distance is smaller than a given threshold $\delta \geq 0$, i.e., condition $h(\mathbf{x}) = h(\mathbf{x}')$ is replaced by $h(\mathbf{x}) \approx_\delta h(\mathbf{x}')$. All exact notions of opacity defined in Definition 3.2.2 can be generalized to the approximate versions by replacing the output equivalence condition as δ-closeness. Note that we will mainly focus on the three basic types of opacity, i.e., initial-state opacity, current-state opacity, and infinite-step opacity.

Definition 3.3.1 (**Approximate Opacity for dt-CS**). Let $\Sigma = (X, X_0, X_S, U, \longrightarrow, Y, h)$ be a metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. System Σ is said to be

- *δ-approximate initial-state opaque* if for any path $\mathbf{x} = x_0 x_1 \cdots x_n \in \mathtt{Path}(\Sigma)$, where $x_0 \in X_S$, there exists path $\mathbf{x}' = x'_0 x'_1 \cdots x'_n \in \mathtt{Path}(\Sigma)$, where $x'_0 \notin X_S$, such that $h(\mathbf{x}) \approx_\delta h(\mathbf{x}')$;
- *δ-approximate current-state opaque* if for any path $\mathbf{x} = x_0 x_1 \cdots x_n \in \mathtt{Path}(\Sigma)$, where $x_n \in X_S$, there exists a path $\mathbf{x}' = x'_0 x'_1 \cdots x'_n \in \mathtt{Path}(\Sigma)$, where $x'_n \notin X_S$, such that $h(\mathbf{x}) \approx_\delta h(\mathbf{x}')$;
- *δ-approximate infinite-step opaque* if for any path $\mathbf{x} = x_0 x_1 \cdots x_n \ldots x_{n+k} \in \mathtt{Path}(\Sigma)$, where $x_n \in X_S$, there is a path $\mathbf{x}' = x'_0 x'_1 \cdots x'_n \ldots x'_{n+k} \in \mathtt{Path}(\Sigma)$, where $x'_n \notin X_S$, such that $h(\mathbf{x}) \approx_\delta h(\mathbf{x}')$.

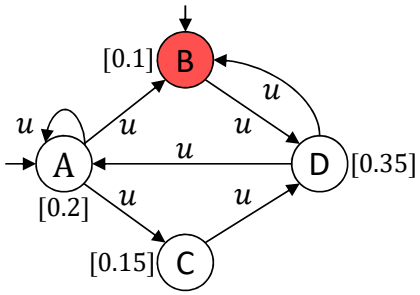


Figure 3.2: An example for approximate opacity, where states marked by red denote secret states, states marked by input arrows denote initial states and the output map is specified by the value associated to each state.

Clearly, when $\delta = 0$, δ-approximate initial-state opacity reduces to its exact version in Definition 3.2.2. The main difference is how we treat two outputs as indistinguishable outputs. Specifically, same as in the exact case, we still assume that the intruder know the system model and the output trajectory generated. However, we further assume that the intruder may not be able to distinguish an output trajectory from other δ-closed ones. Intuitively, the approximate version of opacity can be interpreted as "*the secret of the system cannot be revealed to an intruder that does not have an enough measurement precision related to parameter* δ". In other words, instead of providing an exact security guarantee, approximate opacity provides a relaxed and quantitative security guarantee with respect to the measurement precision of the intruder. Therefore, the value δ can be interpreted as either the measurement imprecision of the intruder or the security level the system can guarantee, i.e., under how powerful intruder the system is still secure.

Hereafter, we assume without loss of generality that

$$\forall x_0 \in X_0 : \{x \in X_0 : \mathbf{d}(h(x_0), h(x)) \leq \delta\} \not\subseteq X_S, \tag{3.3.1}$$

for any system $\Sigma = (X, X_0, X_S, U, \longrightarrow, Y, h)$. This assumption essentially requires that the secret of the system cannot be revealed initially; otherwise, the system is not opaque trivially. This assumption can be easily checked and its non-satisfaction means that δ-approximate initial-state opacity, δ-approximate current-state opacity and δ-approximate infinite-step opacity are all violated trivially.

Example 3.3.2. *Consider system* $\Sigma = (X, X_0, X_S, U, \longrightarrow, Y, h)$ *depicted in Figure 3.2, where* $X = \{A, B, C, D\}, X_0 = \{A, B\}, X_\Sigma = \{B\}, U = \{u\}, h = \{0.1, 0.15, 0.2, 0.35\} \subseteq \mathbb{R}$ *and the output map is specified by the value associated to each state. Clearly, none of exact initial-state opacity, exact current-state opacity and exact infinite-step opacity is satisfied since we know immediately that the system is at secret state* B *when value* 0.1 *is observed.*

Now, let us assume that the output set Y *is equipped with metric* $\mathbf{d}$ *defined by* $\mathbf{d}(y_1, y_2) = |y_1 - y_2|$. *We claim that* S *is not* 0.05*-approximate current-state opaque. For example, let us consider finite run* $B \xrightarrow{u} D \xrightarrow{u} B$ *that generates output run*

$(0.1, 0.35, 0.1)$. *However, there does not exists a finite run leading to a non-secret state whose output run is* 0.05*-close to the above output run. To see this, in order to match the above output run, we must consider a run starting from state* B*, since for the initial state* A*, we have* $\mathbf{d}(h(A), h(B)) = 0.1 \geq 0.05$*, and the next state reached can only be* D*. From state* D*, we can reach states* A *and* B*, but* $\mathbf{d}(h(A), 0.1) = 0.1 \geq 0.05 =: \delta$*. Therefore, the only finite run that approximately matches the above output will end up with secret state* B*, i.e., we know unambiguously that the system is currently at a secret state even when we cannot measure the output precisely. On the other hand, one can check that the system is* 0.1*-approximate current-state opaque. Similarly, system* S *is not* 0.1*-approximate initial-state opaque, since for output run* $(0.1, 0.35)$ *starting from the secret state* B*, there is no run starting from a non-secret initial state that can approximately match it. One can also check that the system is* δ*-approximate initial-state opaque only when* $\delta \geq 0.15$*. We will provide formal procedures for verifying approximate opacity later.*

Remark 3.3.3. *Let* $\Sigma = (X, X_0, X_S, U, \longrightarrow, Y, h)$ *be a metric system. If the output map* h *is identity, i.e.* $h(x) = x$*,* $\forall x \in X$*, then* Σ *is trivially not exactly opaque as in Definition 3.2.2 since we know the exact state of the system directly. However, this is not the case for the approximate notions of opacity as in Definition 3.3.1 since the distance between a secret state and a non-secret state can be very small even if their values are not exactly the same.*

In the sequel, we make some remarks regarding relationships between these notions of state-based opacity. Specifically, we show approximate infinite-step opacity implies approximate initial-state, current-state, and K-step opacity. We should note that similar results does not hold for the lack of opacity.

Lemma 3.3.4. *If a system* $\Sigma = (X, X_0, X_S, U, \longrightarrow, Y, h)$ *is* δ*-approximate infinite-step opaque, it is also* δ*-approximate initial-state opaque (resp. current-state opaque and* K*-step opaque).*

Proof. Let us note that the other notions of approximate opacity as in Definition 3.3.1 can be regarded as special cases of approximate infinite-step opacity. Specifically, as can be seen from Definition 3.3.1, δ-approximate infinite-step opacity requires that the intruder should never know for sure that the system is/was at a secret state for any specific time instant $k \in \{0, \ldots n\}$. When $k = 0$, the notion of approximate infinite-step opacity reduces to approximate initial-state opacity; when $k = n$, infinite-step boils down to current-state opacity. Moreover, note that the notion of K-step opacity requires that the secret should not be revealed within K steps prior to the current instant, while infinite-step opacity captures the entire observation trajectory from initial point up to the current time, which is again stronger then K-step opacity. Therefore, if one can verify that a system Σ is δ-approximate infinite-step opaque, it suffices to claim that Σ is also δ-approximate initial-state opaque (resp. current-state opaque and K-step opaque). □

Remark 3.3.5. *We remark that our notion of initial-state opacity is different from that of* observability*. An observability notion states that every initial state can be*

determined by observing a finite output sequence under a given input run [66]. However, in our context, initial-state opacity is defined as the plausible deniability of a system for every secret initial information under any input sequence. In DESs literature, it was shown that observability can be reformulated as language-based opacity by properly specifying the languages and the observation mapping [107]. However, the relationship between opacity and observability is more challenging in the domain of CPS and is left to future investigation.

3.3.2 Approximate Opacity for Stochastic Control Systems

The notion of approximate opacity proposed in the previous section serves as a foundation for opacity analysis of non-stochastic systems. In particular, the system considered above does not incorporate any stochasticity, and the state runs of the concrete systems are generated non-stochastically under certain input sequences. Whereas in real life applications, many systems are endowed with probabilistic dynamics subject to probabilistic noise and events. Systems with stochastic uncertainties are naturally modelled as general Markov decision processes (gMDPs) as mentioned in Section. 2.3.1.

In this section, we introduce a new notion of initial-state opacity for the class of discrete-time stochastic control systems, called (δ, ε)-approximate initial-state opacity. Our notion can be regarded as the stochastic counterpart of the notion of approximate opacity introduced in Subsection. 3.3.1. In particular, the δ-approximate initial-state opacity proposed in Subsection. 3.3.1 requires that given a threshold parameter $\delta \geq 0$ (based on the measurement precision of the intruder), for any state run starting from a secret state, there always exists another state run starting from a non-secret state, such that the largest distance between their output runs is smaller than δ. In this section, the aim is still to ensure that discerning which of the states was the originating one is difficult for an intruder based on its observation. Particularly, starting from two initial states, the output trajectories are considered indistinguishable if the probability measures of them remaining in any set of interest are close to each other. The parameter ε is used to bound the probability distance and δ captures the measurement precision of the intruder. In the special case when $\delta = 0$, the notion boils down to ε-approximate initial-state opacity, and the parameter ε can be captured *exactly* by the well-known *total variation distance* [36].

In order to show our new notion of opacity, we define for any stochastic control system $\Sigma = (X, X_0, X_S, U, \varsigma, f, Y, h)$, set $B_Y = \{y = (y_0, y_1, \ldots, y_n) \in Y^n | n \in \mathbb{N}_{\geq 1}\}$ which is the set of all finite output sequences. For any measurable set $E \subseteq B_Y$, for some $\delta > 0$, we denote the δ neighborhood of set E by $\underline{E}_\delta$ and $\bar{E}^\delta$, where $\underline{E}_\delta$ is the largest measurable set contained in E satisfying:

$$\underline{E}_\delta = \{y \in E \,|\, \forall \bar{y} \in B_Y \setminus E, \|\bar{y}(i) - y(i)\| \geq \delta, \forall i \in \{0, \ldots, n\}\}, \tag{3.3.2}$$

and $\bar{E}^\delta$ is the smallest measurable set containing E satisfying:

$$\bar{E}^\delta = \{y \in B_Y \,|\, \exists \bar{y} \in E, \|\bar{y}(i) - y(i)\| \leq \delta, \forall i \in \{0, \ldots, n\}\}. \tag{3.3.3}$$

Note that we can regard $\underline{E}_\delta$ and $\bar{E}^\delta$, respectively, as the δ-deflated and δ-inflated version of set E.

Now, we introduce a notion of opacity for the class of stochastic systems defined above.

Definition 3.3.6 (**Approximate Opacity for dt-SCS**). Let $\Sigma = (X, X_0, X_S, U, \varsigma, f, Y, h)$ be a dt-SCS and consider constants $\delta \geq 0$, $0 \leq \varepsilon < 1$. System Σ is (δ, ε)-approximate initial-state opaque if for any $x_0 \in X_0 \cap X_S$, there exists $x_0' \in X_0 \setminus X_S$, so that for any input sequence ν, there exisits an input sequence ν', such that for every measurable set $E \subseteq B_Y$ and the δ neighboring sets $\underline{E}_\delta$ and $\bar{E}^\delta$ as defined in (3.3.2) and (3.3.3), the following inequalities hold:

$$\mathbb{P}(\zeta_{x_0'\nu'} \in \underline{E}_\delta) - \varepsilon \leq \mathbb{P}(\zeta_{x_0\nu} \in E) \leq \mathbb{P}(\zeta_{x_0'\nu'} \in \bar{E}^\delta) + \varepsilon, \tag{3.3.4}$$

where $\zeta_{x_0\nu}$ and $\zeta_{x_0'\nu'}$ are the output runs of the same length, starting from x_0 under ν and starting from x_0' under ν', respectively.

Remark 3.3.7. *In this definition, we use parameter ε to denote the largest allowable probability violation for the output trajectories starting from the secret and non-secret initial states x_0 and x_0' to be δ close. Note that the value of parameter δ is chosen depending on the measurement precision of a malicious intruder. In the case that the precision of the intruder is lower than δ, the δ neighborhood of set E, i.e. $\underline{E}_\delta$ and $\bar{E}^\delta$, is indistinguishable from set E from the intruder's point of view. When $\delta = 0$, the probability inequalities in* (3.3.4) *boils down to $|\mathbb{P}(\zeta_{x_0\nu} \in E) - \mathbb{P}(\zeta_{x_0'\nu'} \in E)| \leq \varepsilon$, and we use the term ε-approximate initial-state opacity. It is worth mentioning that, in this case the parameter ε can be captured exactly by total variation distance [36, 122] for the case of finite MDPs. Thus existing techniques for computing total variation distance can be leveraged as tools to check the probability distance in* (3.3.4)*, which would be applicable for the verification of ε-approximate initial-state opacity for finite MDPs. Although computing this distance is shown to be NP-hard [36, 122], there have been some results to approximate the distance, which are #P-hard and in PSPACE, see e.g. [31, 86]. Moreover, in the case that $\delta = 0$, $\varepsilon = 0$, and no stochasticity exists in the transition functions of the systems, this notion boils down to* exact opacity *as in Definition 3.2.2 defined for finite transition systems.*

3.4 Safety and Security in Formal Methods: Temporal Logic

In the DES literature, opacity is defined over (possibly arbitrarily long) finite paths. In the context of formal verification and synthesis in the computer science literature, formal properties are usually defined over infinite traces. Specifically, a property $\mathcal{P} \subseteq (2^{\mathcal{AP}})^\omega$ is a subset of infinite traces. Since languages over infinite sequences are more expressive than languages over finite ones, it is more general to consider ω-languages than finite-languages. Formal logics such as linear temporal logic (LTL) [13] and their

generalizations (hyperLTL [32, 33]) are convenient ways to express subsets of ω-regular languages.

Safety and Mission Requirements Linear Temporal Logic (LTL) [13] is a convenient and expressive formalism to express properties of infinite runs (or traces) of the system. A restricted form of LTL [40] has been proposed to express properties of finite runs or traces. The set of LTL properties over the atomic proposition $\mathcal{AP}$ can be defined by the following grammar:

$$\phi ::= a \in \mathcal{AP} \mid \neg\phi \mid \phi \vee \phi \mid \mathsf{X}\,\phi \mid \phi\,\mathsf{U}\,\phi.$$

Here, $\neg$ and $\vee$ stand for logical negation and disjunction, while X and U are temporal modalities expressing `next` (in the next discrete step) and `until` (left property continues to hold until the property on the right holds) modalities, respectively. For convenience, additional operators can be derived from these basic ones: $\texttt{true} \stackrel{\text{def}}{=} a \vee \neg a; \texttt{false} \stackrel{\text{def}}{=} \neg\texttt{true}; \varphi \wedge \psi \stackrel{\text{def}}{=} \neg(\neg\varphi \vee \neg\psi); \varphi \rightarrow \psi \stackrel{\text{def}}{=} \neg\varphi \vee \psi; \mathsf{F}\,\varphi \stackrel{\text{def}}{=} \texttt{false}\,\mathsf{U}\,\varphi$; and $\mathsf{G}\,\varphi \stackrel{\text{def}}{=} \neg\,\mathsf{F}\,\neg\varphi$. Here $\wedge$ and $\rightarrow$ stand for conjunction and implication, while F and G stand for temporal operators `finally` (some time in the future) and `globally` (at each step). The semantics of the LTL can be defined inductively in a straightforward fashion (see, [13]). This logic allows the designers to unambiguously characterize system properties. For instance, a safety property can be expressed as "$\mathsf{G}\,\neg\phi$" which states that some bad property ϕ never holds. Similarly, a reachability property "$\mathsf{F}\,\phi$" can be used to express that some good property ϕ eventually holds.

For an infinite trace $r \in \texttt{Trace}^{\omega}(\Sigma)$ of a system Σ, we say that r satisfies the LTL property φ and denoted by $r \models \varphi$, if it satisfies the LTL formula φ. It is known that the set of all infinite traces satisfying an LTL formula can be accepted by either a non-deterministic Büchi automaton or a deterministic Rabin automaton [13]. Given a system Σ and an LTL requirement φ, we denote by $\Sigma \models \varphi$ if for every infinite trace $r \in \texttt{Trace}^{\omega}(\Sigma)$ we have that $r \models \varphi$.

LTL formulae capture the safety and functional correctness requirements of the system. Essentially, it evaluates whether or not each single infinite trace satisfies the property. However, formal reasoning about security properties requires reasoning with multiple traces of the system. For example, Alur et al. [3] show that modal μ-calculus is insufficient to express all opacity policies.

Clarkson and Schneider [33] introduced the concept of hyperproperties to express security policies using second-order logic. Hyperproperties generalize the concept of linear-time properties [13] from being sets of runs to *sets of sets of runs*. HyperLTL, unlike LTL which implicitly considers only a single trace at a time, can relate different trace executions simultaneously through the use of existential and universal quantifiers. The HyperLTL formulae can be given using the following grammar:

$$\begin{aligned} \psi &::= \exists\pi.\psi \mid \forall\pi.\psi \mid \phi \\ \phi &::= a_\pi \mid \neg\phi \mid \phi \vee \phi \mid \mathsf{X}\,\phi \mid \phi\,\mathsf{U}\,\phi. \end{aligned}$$

The key distinction over LTL formulae is the introduction of trace quantifiers $\exists$ and $\forall$. The quantifier $\exists\pi$ stands for "for some trace π" while the quantifier $\forall\pi$ stands for "for all traces π", respectively. The variable ϕ generates standard LTL formulae (complete with Boolean connectives and temporal operators X and U) with the exception that atomic propositions can refer to distinct trace variables. Hence, for every proposition $a \in \mathcal{AP}$ and trace variable π, we use a_π to express that proposition a is referring to the trace π. We say that a trace variable occurs free in a HyperLTL formula, if it is not bounded by any trace quantifier. A HyperLTL formula with no free variable is called a closed formula.

HyperLTL can express certain opacity properties. For instance, the following HyperLTL formula expresses language-based opacity introduced in Definition 3.2.1 when $\mathcal{P}_S$ and $\mathcal{P}_P$ are given as LTL properties ς and φ

$$\forall\pi\exists\pi' \cdot L(\pi) \models \varsigma \rightarrow (h(\pi) = h(\pi') \wedge L(\pi') \models \varphi)$$

where π is defined over $\mathtt{Path}^\omega(\Sigma)$.

Unfortunately, since HyperLTL requires quantification over paths in the beginning of the formula, it is not expressive enough to define infinite-step, current-state, and K-step opacity requirements.

3.5 Generalized Language-Based Opacity

We propose the following generalized language-based opacity notion which extends language-based opacity in Definition 3.2.1 from finite paths to infinite paths.

Definition 3.5.1 (**Generalized Language-Based Opacity**). Let $\Sigma = (X, X_0, U, \longrightarrow, \mathcal{AP}, L, Y, h)$ be a metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$, $\mathcal{P}_S \subseteq \mathtt{Trace}^\omega(\Sigma)$ be a secret property and $\mathcal{P}_P \subseteq \mathtt{Trace}^\omega(\Sigma)$ be a public property. For computational representation, the secret and public properties can be expressed either logically (e.g., via LTL) or using automatic structures (e.g., ω-automata or finite state machines).
We say system Σ is **opaque** with respect to $\mathcal{P}_S$ and $\mathcal{P}_P$ if for any secret path $\mathbf{x} \in \mathtt{Path}^\omega(\Sigma)$, where $L(\mathbf{x}) \in \mathcal{P}_S$, there exists a non-secret path $\mathbf{x}' \in \mathtt{Path}^\omega(\Sigma)$, where $L(\mathbf{x}') \in \mathcal{P}_P$, such that

$$h(\mathbf{x}) \approx_\delta h(\mathbf{x}').$$

The above definition of language-based opacity generalizes Definition 3.2.1 in threefold. First, secret behaviors are defined in terms of traces rather than the internal paths. This setting clearly subsumes Definition 3.2.1 because we can set the labeling function as an identity mapping $L : X \rightarrow X$. Second, secret behaviors are evaluated in terms of infinite sequences rather than finite sequences. Note that, state-based notions of opacity in Definition 3.2.2 are instances of Definition 3.2.1. Therefore, the notions of state-based opacity, such as initial-state opacity or infinite-step opacity, can all be

formulated in terms of Definition 3.5.1 with a syntactic modification to the system (by adding a dummy sink state to the system) to enable the treatment of finite sequences as infinite sequences. Finally, Definition 3.5.1 considers approximate output equivalence rather than the exact one. Language-based opacity in Definition 3.5.1 also generalizes the notions of *noninterference* [128, 134, 201] and 2-safety [178, 139].

3.6 Discussion

This chapter provides the basic foundations on which related verification and synthesis questions can be posed and answered. We developed a unifying common framework which integrates ideas and formalism from three distinct fields of formal methods, discrete-event systems, and control theory. We reviewed the typical security notions in both discrete-event systems and formal methods communities, and then, proposed novel concepts of security notions, called approximate opacity, which are more suitable for (stochastic) CPS. The security notions and mission requirements from different fields were then unified as a generalized language-based opacity notion.

4 Abstraction-based Opacity Verification of Cyber-physical Systems

4.1 Introduction

Cyber-physical systems (CPS) are complex systems resulting from tight interactions of dynamical systems and computational devices. Models of CPS are inherently heterogeneous: from discrete systems modeling computational parts to differential or difference equations modeling continuous physical processes. Such heterogeneity makes the verification and design of such systems significantly challenging. In addition, components in CPS are usually connected via communication networks in order to acquire and exchange information so that some global functionality of the system can be achieved. However, this also brings new challenges for the verification and design of CPS since the communication between system components may release information that might compromise the security of the system. Therefore, how to analyze and enforce security for CPS is becoming an increasingly important issue and has drawn considerable attention in the literature in the past few years [90, 167].

In order to address the heterogeneity of CPS models, formal verification and synthesis are often addressed by methods of *abstraction* in which continuous-space models are approximated by discrete ones. When a suitable finite abstraction is constructed, by leveraging computational tools developed for DES and games on automata, one can verify or synthesize controllers in an automated fashion against complex logic requirements. The pipeline of traditional abstraction-based verification technique is depicted in Figure 4.1, which consists of three key phases. The first phase is on the construction of a finite abstraction of the CPS with the property that the set of behaviours of the CPS is included in that of the constructed finite abstraction. The second phase in the architecture requires symbolic analysis to efficiently reason about formal specifications. The final phase is to bring the reasoning back to the original concrete systems with formal guarantee. The key to the construction of such finite/symbolic systems is the establishment of formal relations between the concrete and abstract systems. A system relation formalizes the ability to extrapolate properties from an abstraction to the concrete system. Different system relations enable extrapolation of different kinds of properties. Such relations include (alternating) (bi)simulation relations, their approximate versions, and strongest or asynchronous ℓ-complete approximations.

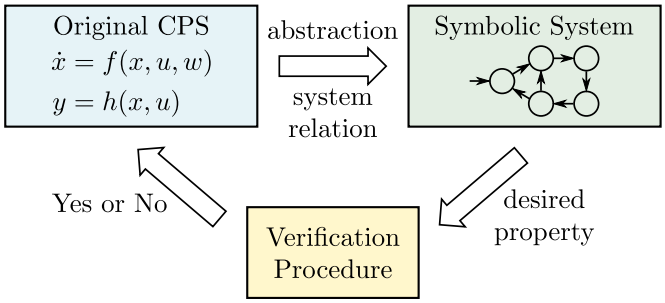


Figure 4.1: Pipeline of standard discretization-based or abstraction-based verification technique.

4.1.1 Related Literature

Finite abstraction together with the notions of so-called simulation relations have been widely and successfully used in the past decade for formal verification, synthesis, and approximation of hybrid systems [4, 51, 50, 53, 218, 214, 145, 186, 17, 216, 157]. Nevertheless, none of the constructed finite abstractions in the aforementioned literature is guaranteed to preserve opacity. As reported in [222], existing notions of standard (bi)simulation relations and their approximate versions which are often used in finite abstraction synthesis schemes fail to preserve opacity.

In the recent literature, there have been several attempts on leveraging abstraction-based techniques for the verification and synthesis of opacity; see, e.g., [222, 135, 136, 199, 131, 69]. In particular, in a recent work [222], the authors propose several notions of opacity-preserving (bi)simulation relations. However, these relations only preserve exact opacity for non-metric systems. Motivated by this, we will provide new relations to extend the relations in [222] to metric systems by taking into account how close two systems are. We need to consider both the dynamic and the secret of the system while constructing the symbolic model and guarantee the preservation of approximate opacity across related systems.

On the other hand, in many real-world applications, a small probability of violation of the opacity could be tolerable. Hence, instead of simply asking if a system is opaque or non-opaque, it is more applicable to evaluate the possibility of being not secure for stochastic systems. This direction has been recently explored in the context of stochastic DES [163, 18, 30, 2, 199, 211]. For example, in [163] three different notions of probabilistic opacity were introduced for current-state opacity; this approach has also been extended to infinite-step opacity by [211]. In [30], Jensen-Shannon divergence was adopted to quantify secrecy loss in stochastic systems. Opacity of (partially-observed) Markov decision processes has also been studied in [18, 2, 199]. Note that most of the existing works on opacity analysis of stochastic DES are based on finite systems. In discrete-time stochastic control systems, however, the state sets are usually infinite, which makes the verification problem very challenging and even undecidable. Therefore, efficient abstraction techniques, together with suitably defined notions of stochastic

opacity, are needed in order to quantitatively evaluate the security level of continuous-space stochastic control systems.

4.1.2 Contributions

In the previous chapter, we introduced new concepts of approximate opacity suitable for (stochastic) control systems. In this chapter, we focus on verifying approximate opacity for both non-stochastic and stochastic control systems. In the context of (non-stochastic) discrete-time control systems, we propose a new simulation-type relation, called approximate opacity-preserving simulation relation, which characterizes how close two systems are in terms of the satisfaction of approximate opacity. This allows us to verify approximate opacity for large-scale, or even infinite systems, using their abstractions. Effective algorithms are provided to verify different notions of approximate opacity. Then, for a class of incrementally input-to-state stable discrete-time control systems with possibly continuous state sets, we propose an effective approach to construct symbolic models (a.k.a. finite abstractions) that approximately simulate the original systems in the sense of opacity preserving and vice versa. Therefore, the proposed abstraction technique together with the verification algorithm for the finite case provide a sound way for verifying opacity of discrete-time control systems with continuous state sets. In the context of discrete-time stochastic control systems, we introduce a new notion of so-called initial-state opacity-preserving stochastic simulation functions to quantify the distance between two systems in a probabilistic setting, while preserving approximate initial-state opacity across them. This allows us to efficiently verify opacity of a complex stochastic system with possibly an uncountable state set by analyzing it over its simpler (potentially finite) abstraction. In addition, we show that for a class of stochastic control systems satisfying incremental input-to-state stability property, one can construct their finite abstractions (a.k.a finite Markov decision processes (MDP)) together with a corresponding opacity-preserving stochastic simulation function between them.

4.2 Verification of Approximate Opacity for Finite Systems

In this section, we first show how to verify the three basic types of approximate opacity (i.e., approximate initial-state opacity, approximate current-state opacity, and approximate infinite-step opacity) for finite systems. This will provide the basis for the verification of approximate opacity for infinite systems.

4.2.1 Verification of Approximate Initial-State Opacity

In order to verify δ-approximate initial-state opacity as in Definition 3.3.1, we construct a new system called the *δ-approximate initial-state estimator* defined as follows.

Definition 4.2.1. Let $\Sigma = (X, X_0, X_S, U, \longrightarrow, Y, h)$ be a metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. The δ-approximate initial-state estimator is a system (without outputs)

$$\Sigma_I = (X_I, X_{I0}, U, \xrightarrow[I]{}),$$

where

- $X_I \subseteq X \times 2^X$ is the set of states;
- $X_{I0} = \{(x, q) \in X \times 2^X : x' \in q \Leftrightarrow \mathbf{d}(h(x), h(x')) \leq \delta\}$ is the set of initial states;
- U is the set of inputs, which is the same as the one in Σ;
- $\xrightarrow[I]{} \subseteq X_I \times U \times X_I$ is the transition function defined by: for any $(x, q), (x', q') \in X \times 2^X$ and $u \in U$, $(x, q) \xrightarrow[I]{u} (x', q')$ if
  1. $(x', u, x) \in \longrightarrow$; and
  2. $q' = \cup_{\hat{u} \in U} \mathbf{Pre}_{\hat{u}}(q) \cap \{x'' \in X : \mathbf{d}(h(x'), h(x'')) \leq \delta\}$.

For the sake of simplicity, we only consider the part of Σ_I that is reachable from initial states.

Intuitively, the δ-approximate initial-state estimator works as follows. Each initial state of Σ_I is a pair consisting of a system state and its δ-closed states; we consider all each pairs as the set of initial states. Then from each state, we track *backwards* states that are consistent with the output information recursively. Our construction is motivated by the reversed-automaton-based initial-state-estimator proposed in [202] but with the following differences. First, the way we defined information-consistency is different. Here we treat states whose output are δ-close to each other as consistent states. Moreover, the structure in [202] only requires a state space of 2^X, while our state space is $X \times 2^X$. The additional first component can be understood as the "reference trajectory" that is used to determine what is "δ-close" at each instant. We use the following result to show the main property of Σ_I.

Proposition 4.2.2. *Let $\Sigma = (X, X_0, X_S, U, \xrightarrow[I]{}, Y, h)$ be a metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. Let $\Sigma_I = (X_I, X_{I0}, U, \xrightarrow[I]{})$ be its δ-approximate initial-state estimator. Then for any $(x_0, q_0) \in X_{I0}$ and any finite run*

$$(x_0, q_0) \xrightarrow[I]{u_1} (x_1, q_1) \xrightarrow[I]{u_2} \cdots \xrightarrow[I]{u_n} (x_n, q_n),$$

we have

(i) $x_n \xrightarrow{u_n} x_{n-1} \xrightarrow{u_{n-1}} \cdots \xrightarrow{u_1} x_0$; *and*

$$\textit{(ii)}\ q_n = \left\{ x_0' \in X : \begin{array}{c} \exists x_0' \xrightarrow{u_n'} x_1' \xrightarrow{u_{n-1}'} \cdots \xrightarrow{u_1'} x_n' \ s.t. \\ \max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x_{n-i}')) \leq \delta \end{array} \right\}.$$

Proof. It is straightforward to show (i). Hereafter, we prove (ii) by induction on the length of input sequence.

When $n = 0$, i.e., there is no input sequence, we have that $(x_0, q_0) \in X_{I0}$. By the definition of X_{I0}, we know that

$$q_0 = \{x_0' \in X : \mathbf{d}(h(x_0), h(x_0')) \leq \delta\},$$

which implies (ii) immediately.

To proceed the induction, we assume that (ii) holds when $n = k$. Now, we need to show that (ii) also holds when $n = k + 1$. Consider arbitrary pair $(x_0, q_0) \in X_{I0}$ and finite run

$$(x_0, q_0) \xrightarrow[I]{u_1} (x_1, q_1) \xrightarrow[I]{u_2} \cdots \xrightarrow[I]{u_n} (x_n, q_n) \xrightarrow[I]{u_{n+1}} (x_{n+1}, q_{n+1}).$$

Then, we have

$$\begin{aligned} q_{n+1} =& \cup_{\hat{u}\in U} \mathbf{Pre}_{\hat{u}}(q_n) \cap \{x \in X : \mathbf{d}(h(x_{n+1}), h(x)) \leq \delta\} \\ =& \{x \in X : \exists x' \in q_n, u_{n+1}' \in U \text{ s.t. } (x, u_{n+1}', x') \in \longrightarrow \} \\ & \cap \{x \in X : \mathbf{d}(h(x_{n+1}), h(x)) \leq \delta\} \\ =& \left\{ x \in X : \begin{array}{c} [\exists x' \in q_n, u_{n+1}' \in U \text{ s.t. } (x, u_{n+1}', x') \in \longrightarrow] \\ \wedge [\mathbf{d}(h(x_{n+1}), h(x)) \leq \delta] \end{array} \right\}. \end{aligned}$$

By the induction hypothesis, we know that

$$q_n = \left\{ x_0' \in X : \begin{array}{c} \exists x_0' \xrightarrow{u_n'} x_1' \xrightarrow{u_{n-1}'} \cdots \xrightarrow{u_1'} x_n' \text{ s.t.} \\ \max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x_{n-i}')) \leq \delta \end{array} \right\}.$$

Therefore, by combing the above two equations, one gets

$$\begin{aligned} q_{n+1} =& \left\{ x \in X : \begin{array}{c} \exists x \xrightarrow{u_{n+1}'} x_0' \xrightarrow{u_n'} x_1' \xrightarrow{u_{n-1}} \cdots \xrightarrow{u_1'} x_n' \\ \text{s.t. } \max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x_{n-i}')) \leq \delta \\ \wedge \mathbf{d}(h(x_{n+1}), h(x)) \leq \delta \end{array} \right\} \\ =& \left\{ x \in X : \begin{array}{c} \exists x_0'' \xrightarrow{u_{n+1}'} x_1'' \xrightarrow{u_n'} \cdots \xrightarrow{u_1'} x_{n+1}'' \text{ s.t.} \\ \max_{i\in\{0,1,\ldots,n+1\}} \mathbf{d}(h(x_i), h(x_{n+1-i}'')) \leq \delta \end{array} \right\}. \end{aligned}$$

Therefore, one obtains that the induction step holds. □

The next theorem provides one of the main results of this section on the verification of δ-approximate initial-state opacity of finite metric systems.

Theorem 4.2.3. Let $\Sigma = (X, X_0, X_S, U, \longrightarrow, Y, h)$ be a finite metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. Let $\Sigma_I = (X_I, X_{I0}, U, \underset{I}{\longrightarrow})$ be its δ-approximate initial-state estimator. Then, Σ is δ-approximate initial-state opaque if and only if

$$\forall (x, q) \in X_I : x \in X_0 \cap X_S \Rightarrow q \cap X_0 \not\subseteq X_S. \tag{4.2.1}$$

Proof. ($\Rightarrow$) By contraposition: suppose that there exists a state $(x, q) \in X_I$ such that $x \in X_0 \cap X_S$ and $q \cap X_0 \subseteq X_S$. Let

$$(x_0, q_0) \underset{I}{\longrightarrow} (x_1, q_1) \underset{I}{\longrightarrow} \cdots \underset{I}{\longrightarrow} (x_n, q_n),$$

be a run reaching $(x, q) =: (x_n, q_n)$. By Proposition 4.2.2, we have $x_n \xrightarrow{u_n} x_{n-1} \xrightarrow{u_{n-1}} \cdots \xrightarrow{u_1} x_1$, which is well-defined in Σ as $x_n \in X_0$. Moreover, by Proposition 4.2.2, we have

$$q_n = \left\{ x'_0 \in X : \begin{array}{c} \exists x'_0 \xrightarrow{u'_n} x'_1 \xrightarrow{u'_{n-1}} \cdots \xrightarrow{u'_1} x'_n \text{ s.t.} \\ \max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x'_{n-i})) \leq \delta \end{array} \right\}.$$

However, since $q_n \cap X_0 \subseteq X_S$, we know that there does not exist $x'_0 \in X_0 \setminus X_S$ and $x'_0 \xrightarrow{u'_n} x'_1 \xrightarrow{u'_{n-1}} \cdots \xrightarrow{u'_1} x'_n$ such that $\max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x'_{n-i})) \leq \delta$. Therefore, by considering $x_n \in X_0 \cap X_S$ and $x_n \xrightarrow{u_n} x_{n-1} \xrightarrow{u_{n-1}} \cdots \xrightarrow{u_1} x_1$, we know the system is not δ-approximate initial-state opaque.

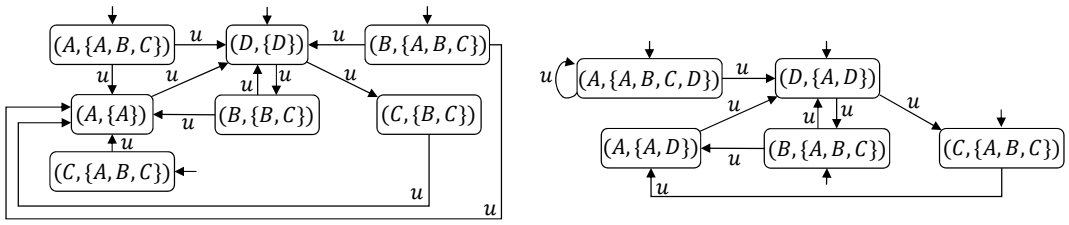
($\Leftarrow$) By contradiction: suppose that Equation (4.2.1) holds and assume that Σ is not δ-approximate initial-state opaque. Then, there exists a secret initial state $x_0 \in X_0 \cap X_S$ and a sequence of transitions $x_0 \xrightarrow{u_1} x_1 \xrightarrow{u_2} \cdots \xrightarrow{u_n} x_n$ such that there does not exist a non-secret initial state $x'_0 \in X_0 \setminus X_S$ and a sequence of transitions $x'_0 \xrightarrow{u'_1} x'_1 \xrightarrow{u'_2} \cdots \xrightarrow{u'_n} x'_n$ such that $\max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x'_i)) \leq \delta$. Let us consider the following sequence of transitions in Σ_I

$$(x_n, q_0) \underset{I}{\xrightarrow{u_n}} (x_{n-1}, q_1) \underset{I}{\xrightarrow{u_{n-1}}} \cdots \underset{I}{\xrightarrow{u_1}} (x_0, q_n).$$

By Proposition 4.2.2, we know that

$$q_n = \left\{ x'_0 \in X : \begin{array}{c} \exists x'_0 \xrightarrow{u'_n} x'_1 \xrightarrow{u'_{n-1}} \cdots \xrightarrow{u'_1} x'_n \text{ s.t.} \\ \max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x'_i)) \leq \delta \end{array} \right\}.$$

By Equation (4.2.1), we have $q_n \cap X_0 \not\subseteq X_S$. Therefore, there exist a non-secret initial state $x'_0 \in X_0 \setminus X_S$ and a sequence $x'_0 \xrightarrow{u'_1} x'_1 \xrightarrow{u'_2} \cdots \xrightarrow{u'_n} x'_n$ such that $\max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x'_i)) \leq \delta$. This is a contradiction, i.e., Σ has to be δ-approximate initial-state opaque. □



(a) S_I when $\delta = 0.1$ (b) S_I when $\delta = 0.15$

Figure 4.2: Examples of δ-approximate initial-state estimators.

We illustrate how to verify δ-approximate initial-state opacity by the following example.

Example 4.2.4. *Let us still consider system Σ shown in Figure 3.2. The δ-approximate initial-state estimator Σ_I when $\delta = 0.1$ is shown in Figure 4.2(a). For example, for initial state $(D, \{D\})$, we have $(D, \{D\}) \xrightarrow[I]{u} (B, \{B, C\})$ since $B \xrightarrow{u} D$ and $\{B, C\} = \mathbf{Pre}_u(\{D\}) \cap \{x \in X : \mathbf{d}(0.1, h(x)) \leq 0.1\} = \{B, C\} \cap \{A, B, C\}$. However, for state $(B, \{B, C\}) \in X_I$, we have $B \in X_0 \cap X_S$ and $\{B, C\} \cap X_0 = \{B\} \subseteq X_S$. Therefore, by Theorem 4.2.3, we know that the system is not 0.1-approximate initial-state opaque. Similarly, we can also construct Σ_I for the case of $\delta = 0.15$, which is shown in Figure 4.2(b). Since for state $(B, \{A, B, C\}) \in X_I$, which is the only state whose first component is in $X_0 \cap X_S$, we have $\{A, B, C\} \cap X_0 = \{A, B\} \not\subseteq X_S$. By Theorem 4.2.3, we know that the system is 0.15-approximate initial-state opaque.*

4.2.2 Verification of Approximate Current-State Opacity

In order to verify δ-approximate current-state opacity, we also need to construct a new system called the *δ-approximate current-state estimator* defined as follows.

Definition 4.2.5. *Let $\Sigma = (X, X_0, X_S, U, \longrightarrow, Y, h)$ be a metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. The δ-approximate current-state estimator is a system (without outputs)*

$$\Sigma_C = (X_C, X_{C0}, U, \xrightarrow[C]{}),$$

where

- *$X_C \subseteq X \times 2^X$ is the set of states;*
- *$X_{C0} = \{(x, q) \in X_0 \times 2^{X_0} : x' \in q \Leftrightarrow \mathbf{d}(h(x), h(x')) \leq \delta\}$ is the set of initial states;*
- *U is the set of inputs, which is the same as the one in Σ;*

- $\xrightarrow[C]{} \subseteq X_C \times U \times X_C$ *is the transition function defined by: for any* $(x,q),(x',q') \in X \times 2^X$ *and* $u \in U$, $(x,q) \xrightarrow[C]{u} (x',q')$ *if*
  1. $(x,u,x') \in \longrightarrow$ *; and*
  2. $q' = \cup_{\hat{u} \in U} \mathbf{Post}_{\hat{u}}(x) \cap \{x'' \in X : \mathbf{d}(h(x'),h(x'')) \leq \delta\}$.

For the sake of simplicity, we only consider the part of Σ_C *that is reachable from initial states.*

The construction of Σ_C is similar to Σ_I. However, we need to track all forward runs from each pair of initial-state and its information-consistent states. Still, we need the first component as the "reference state" to determine what are "δ-close" states. We use the following result to state the main properties of Σ_C.

Proposition 4.2.6. *Let* $\Sigma = (X, X_0, X_S, U, \longrightarrow, Y, h)$ *be a metric system, with the metric* $\mathbf{d}$ *defined over the output set, and a constant* $\delta \geq 0$. *Let* $\Sigma_C = (X_C, X_{C0}, U, \xrightarrow[C]{})$ *be its* δ*-approximate current-state estimator. Then for any* $(x_0,q_0) \in X_{C0}$ *and any finite run*

$$(x_0,q_0) \xrightarrow[C]{u_1} (x_1,q_1) \xrightarrow[C]{u_2} \cdots \xrightarrow[C]{u_n} (x_n,q_n),$$

we have

(i) $x_0 \xrightarrow{u_1} x_1 \xrightarrow{u_2} \cdots \xrightarrow{u_n} x_n$*; and*

(ii) $q_n = \{x'_n \in X : \exists x'_0 \in X_0, \exists x'_0 \xrightarrow{u'_1} x'_1 \xrightarrow{u'_2} \cdots \xrightarrow{u'_n} x'_n \text{ s.t. } \max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(h(x_i),h(x'_i)) \leq \delta\}$.

Proof. The proof is similar to that of Proposition 4.2.2, which can be done by induction on the length of the sequence. □

Now, we show the second main result of this section by providing a verification scheme for δ-approximate current-state opacity of finite metric systems.

Theorem 4.2.7. Let $\Sigma = (X, X_0, X_S, U, \longrightarrow, Y, h)$ be a metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. Let $\Sigma_C = (X_C, X_{C0}, U, \xrightarrow[C]{})$ be its δ-approximate current-state estimator. Then, Σ is δ-approximate current-state opaque if and only if

$$\forall (x,q) \in X_C : q \not\subseteq X_S. \tag{4.2.2}$$

Proof. By Proposition 4.2.6, for each state (x,q) encountered, the second component is exactly the set of all possible current states consistent with the observation. Then the proof is similar to that of Theorem 4.2.3. □

4.2.3 Verification of Approximate Infinite-Step Opacity

Finally, we combine the δ-approximate initial-state estimator Σ_I and the δ-approximate current-state estimator Σ_C to verify δ-approximate infinite-step opacity of finite metric systems. The verification scheme is provided by the following theorem.

Theorem 4.2.8. Let $\Sigma = (X, X_0, X_S, U, \longrightarrow, Y, h)$ be a finite metric system, with the metric $\mathbf{d}$ defined over the output set, and a constant $\delta \geq 0$. Let $\Sigma_I = (X_I, X_{I0}, U, \xrightarrow[I]{})$ and $\Sigma_C = (X_C, X_{C0}, U, \xrightarrow[C]{})$ be its δ-approximate initial-state estimator and δ-approximate current-state estimator, respectively. Then, Σ is δ-approximate infinite-step opaque if and only if

$$\forall (x, q) \in X_I, (x', q') \in X_C : x = x' \in X_S \Rightarrow q \cap q' \not\subseteq X_S. \tag{4.2.3}$$

Proof. By contraposition: suppose that there exist two states $(x_n, q'_n) \in X_I, (x_n, q_n) \in X_C$ such that $x_n \in X_S$ and $q_n \cap q'_n \subseteq X_S$. Let

$$(x_0, q_0) \xrightarrow[C]{u_1} (x_1, q_1) \xrightarrow[C]{u_2} \cdots \xrightarrow[C]{u_n} (x_n, q_n),$$
$$(x_{n+m}, q_{n+m}) \xrightarrow[I]{u_{n+m}} (x_{n+m-1}, q_{n+m-1}) \cdots \xrightarrow[I]{u_{n+1}} (x_n, q'_n),$$

be two runs reaching (x, q) and (x, q'), respectively. By Propositions 4.2.2 and 4.2.6, we have $x_0 \in X_0$ and

$$x_0 \xrightarrow{u_1} \cdots \xrightarrow{u_{n-1}} x_{n-1} \xrightarrow{u_n} x_n \xrightarrow{u_{n+1}} x_{n+1} \xrightarrow{u_{n+2}} \cdots \xrightarrow{u_{n+m}} x_{n+m}.$$

Moreover, one has

$$q_n \cap q'_n = \left\{ x'_n \in X : \begin{array}{c} \exists x'_0 \in X_0, \exists x'_0 \xrightarrow{u'_1} \cdots \xrightarrow{u'_{n+m}} x'_{n+m} \\ \text{s.t. } \max_{i \in \{0,1,\ldots,n+m\}} \mathbf{d}(h(x_i), h(x'_i)) \leq \delta \end{array} \right\}.$$

Since $q_n \cap q'_n \subseteq X_S$, we know that there does not exist $x'_0 \in X_0$ and $x'_0 \xrightarrow{u'_1} \cdots \xrightarrow{u'_{n+m}} x'_{n+m}$ such that $x'_n \in X \setminus X_S$ and $\max_{i \in \{0,1,\ldots,n+m\}} \mathbf{d}(h(x_i), h(x'_i)) \leq \delta$. Therefore, the system is not δ-approximate infinite-step opaque.

($\Leftarrow$) By contradiction: suppose that equation (4.2.3) holds and assume, for the sake of contradiction, that Σ is not δ-approximate infinite-step opaque. Then, we know that there exists an initial state $x_0 \in X_0$, a sequence of transitions $x_0 \xrightarrow{u_1} x_1 \xrightarrow{u_2} \cdots \xrightarrow{u_n} x_n$ and an index $k \in \{0, \ldots, n\}$ such that $x_k \in X_S$ and there does not exist an initial state $x'_0 \in X_0$ and a sequence of transitions $x'_0 \xrightarrow{u'_1} x'_1 \xrightarrow{u'_2} \cdots \xrightarrow{u'_n} x'_n$ such that $x'_k \in X \setminus X_S$ and $\max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x'_i)) \leq \delta$. Let us consider the following sequence of transitions in Σ_C

$$(x_0, q_0) \xrightarrow[C]{u_1} (x_1, q_1) \xrightarrow[C]{u_2} \cdots \xrightarrow[C]{u_k} (x_k, q_k),$$

and the following sequence of transitions in Σ_I

$$(x_n, q'_n) \xrightarrow[I]{u_n} (x_{n-1}, q'_{n-1}) \xrightarrow[I]{u_{n-1}} \cdots \xrightarrow[I]{u_{k+1}} (x_k, q'_k).$$

By Propositions 4.2.2 and 4.2.6, we know that

$$q_n \cap q'_n = \left\{ x'_k \in X : \begin{array}{c} \exists x'_0 \in X_0, \exists x'_0 \xrightarrow{u'_1} \cdots \xrightarrow{u'_n} x'_n \text{ s.t.} \\ \max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x'_i)) \leq \delta \end{array} \right\}.$$

Since equation (4.2.3) holds, we know that $q_n \cap q'_n \not\subseteq X_S$. Therefore, there exists $x'_0 \in X_0$ and a sequence of transitions $x'_0 \xrightarrow{u'_1} \cdots \xrightarrow{u'_n} x'_n$ such that $x_k \in X \setminus X_S$ and $\max_{i\in\{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x'_i)) \leq \delta$, which is a contradiction, i.e., Σ has to be δ-approximate infinite-step opaque. □

Remark 4.2.9. *We conclude this section by discussing the complexity of verifying approximate opacity. Let $\Sigma = (X, X_0, X_S, U, \longrightarrow, Y, h)$ be a finite metric system. The complexity of the verification algorithms for both approximate initial-state and current-state opacity is $O(|U| \times |X| \times 2^{|X|})$, which is the size of Σ_I or Σ_C. For approximate infinite-step opacity, we need to construct both Σ_I and Σ_C, and compare each pair of states in Σ_I and Σ_C. Therefore, the complexity for verifying approximate infinite-step opacity using Theorem 4.2.8 is $O(|U| \times |X|^2 \times 4^{|X|})$. It is worth noting that the complexity of verifying exact opacity as in Definition 3.2.2 is already known to be PSPACE-complete [27]. Using a similar reduction, we can conclude that the complexity of verifying approximate opacity as in Definition 3.3.1 is also PSPACE-complete for $\delta > 0$. Finally, we note that the exponential complexity essentially comes from the subset construction to handle information uncertainty. In practice, the subset construction usually results in a quite small structure; see, e.g., [34] for detailed empirical studies on this issue.*

4.3 Approximate Simulation Relations for Opacity

In the previous sections, we introduced the verification procedures of various notions of approximate opacity. However, when the system is very large or even infinite, verifying opacity based on the original system is not efficient or not even possible. Therefore, it will be beneficial if we can verify opacity based on an "equivalent" smaller or symbolic system. To this end, in this section, we study under what conditions two systems are equivalent and in what sense. Specifically, we introduce new notions of approximate opacity-preserving simulation relations, inspired by the one in [51]. The newly proposed simulation relations will provide the basis for abstraction-based verification of approximate opacity.

4.3.1 Approximate Initial-State opacity-preserving Simulation Relation

First, we introduce a new notion of approximate initial-state opacity-preserving simulation relation.

Definition 4.3.1. *(Approximate Initial-State opacity-preserving Simulation Relation) Consider two metric systems* $\Sigma = (X, X_0, X_S, U, f, Y, h)$ *and* $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h})$ *with the same output sets* $Y = \hat{Y}$ *and metric* $\mathbf{d}$*. For* $\varepsilon \in \mathbb{R}_{\geq 0}$*, a relation* $R \subseteq X \times \hat{X}$ *is called an* ε*-approximate initial-state opacity-preserving simulation relation (*ε*-InitSOP simulation relation) from* Σ *to* $\hat{\Sigma}$ *if*

1. *a)* $\forall x_0 \in X_0 \cap X_S, \exists \hat{x}_0 \in \hat{X}_0 \cap \hat{X}_S : (x_0, \hat{x}_0) \in R$*;*
   b) $\forall \hat{x}_0 \in \hat{X}_0 \setminus \hat{X}_S, \exists x_0 \in X_0 \setminus X_S : (x_0, \hat{x}_0) \in R$*;*
2. $\forall (x, \hat{x}) \in R : \mathbf{d}(h(x), \hat{h}(\hat{x})) \leq \varepsilon$*;*
3. *For any* $(x, \hat{x}) \in R$*, we have*
   a) $\forall x \xrightarrow{u} x', \exists \hat{x} \xrightarrow{\hat{u}} \hat{x}' : (x', \hat{x}') \in R$*;*
   b) $\forall \hat{x} \xrightarrow{\hat{u}} \hat{x}', \exists x \xrightarrow{u} x' : (x', \hat{x}') \in R$*.*

We say that Σ *is* ε*-InitSOP simulated by* $\hat{\Sigma}$*, denoted by* $\Sigma \preceq_I^{\varepsilon} \hat{\Sigma}$*, if there exists an* ε*-InitSOP simulation relation* R *from* Σ *to* $\hat{\Sigma}$*.*

Note that a system $\hat{\Sigma}$ that simulates Σ through the InitSOP simulation relation is often called an opacity-preserving abstraction of Σ. We should mention that, although the above relation appears to be similar to the approximate bisimulation relation proposed in [51], it is still a one-sided relation here because Condition 1 is not symmetric. We refer the interested readers to [222] to see why one needs the strong condition 3 in Definition 4.3.1 to show preservation of initial-state opacity in one direction when $\varepsilon = 0$.

The following main theorem provides a sufficient condition for δ-approximate initial-state opacity based on related systems as in Definition 4.3.1.

Theorem 4.3.2. Consider two metric systems $\Sigma = (X, X_0, X_S, U, f, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h})$ with the same output sets $Y = \hat{Y}$ and metric $\mathbf{d}$ and let $\varepsilon, \delta \in R_{\geq 0}$. If $\Sigma \preceq_I^{\varepsilon} \hat{\Sigma}$ and $\varepsilon \leq \frac{\delta}{2}$, then we have:

$$\begin{aligned} &\hat{\Sigma} \text{ is } (\delta - 2\varepsilon)\text{-approximate initial-state opaque} \\ \Rightarrow &\Sigma \text{ is } \delta\text{-approximate initial-state opaque.} \end{aligned}$$

Proof. Consider an arbitrary secret initial state $x_0 \in X_0 \cap X_S$ and a run $\mathbf{x} = x_0 x_1 \cdots x_n$ in Σ. Since $\Sigma \preceq_I^{\varepsilon} \hat{\Sigma}$, by conditions 1a), 2 and 3a) in Definition 4.3.1, there exist a secret initial state $\hat{x}_0 \in \hat{X}_0 \cap \hat{X}_S$ and a run $\hat{\mathbf{x}} = \hat{x}_0 \hat{x}_1 \cdots \hat{x}_n$ in $\hat{\Sigma}$ such that

$$\forall i \in \{0, 1, \ldots, n\} : \mathbf{d}(h(x_i), \hat{h}(\hat{x}_i)) \leq \varepsilon. \tag{4.3.1}$$

Since $\hat{\Sigma}$ is $(\delta - 2\varepsilon)$-approximate initial-state opaque, there exist a non-secret initial state $\hat{x}'_0 \in \hat{X}_0 \setminus \hat{X}_S$ and a run $\hat{\mathbf{x}}' = \hat{x}'_0 \hat{x}'_1 \cdots \hat{x}'_n$ such that

$$\max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(\hat{h}(\hat{x}_i), \hat{h}(\hat{x}'_i)) \leq \delta - 2\varepsilon. \tag{4.3.2}$$

Again, since $\Sigma \preceq_I^{\varepsilon} \hat{\Sigma}$, by conditions 1b), 2 and 3b) in Definition 4.3.1, there exist an initial state $x_0' \in X_0 \setminus X_S$ and a run $\mathbf{x}' = x_0' x_1' \cdots x_n'$ such that

$$\forall i \in \{0, 1, \ldots, n\} : \mathbf{d}(h(x_i'), \hat{h}(\hat{x}_i')) \leq \varepsilon. \tag{4.3.3}$$

Combining equations (4.3.1), (4.3.2), (4.3.3), and using the triangle inequality, we have

$$\max_{i \in \{0,1,\ldots,n\}} : \mathbf{d}(h(x_i), h(x_i')) \leq \delta. \tag{4.3.4}$$

Since $x_0 \in X_0 \cap X_S$ and the run $\mathbf{x} = x_0 x_1 \cdots x_n$ are arbitrary, we conclude that Σ is δ-approximate initial-state opaque. □

The following corollary is a simple consequence of the result in Theorem 4.3.2 but for the lack of δ-approximate initial-state opacity.

Corollary 4.3.3. *Consider two metric systems $\Sigma = (X, X_0, X_S, U, f, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h})$ with the same output sets $Y = \hat{Y}$ and metric $\mathbf{d}$ and let $\varepsilon, \delta \in \mathbb{R}_{\geq 0}$. If $\hat{\Sigma} \preceq_I^{\varepsilon} \Sigma$, then the following implication hold:*

$$\begin{aligned} &\hat{\Sigma} \textit{ is not } (\delta + 2\varepsilon)\textit{-approximate initial-state opaque} \\ \Rightarrow &\Sigma \textit{ is not } \delta\textit{-approximate initial-state opaque.} \end{aligned}$$

Proof. Since $\hat{\Sigma} \preceq_I^{\varepsilon} \Sigma$, by Theorem 4.3.2, we know that Σ being δ-approximate initial-state opaque implies that $\hat{\Sigma}$ is $(\delta + 2\varepsilon)$-approximate initial-state opaque. Hence, $\hat{\Sigma}$ not being $(\delta + 2\varepsilon)$-approximate initial-state opaque implies that Σ is not δ-approximate initial-state opaque. □

Remark 4.3.4. *It is worth remarking that δ and ε are parameters specifying two different types of precision. Parameter δ is used to specify the measurement precision under which we can guarantee opacity for a single system, while parameter ε is used to characterize the "distance" between two systems in terms of being approximate opaque. The reader should not be confused by the different roles of these two parameters.*

We illustrate ε-approximate initial-state opacity-preserving simulation relation by the following example.

Example 4.3.5. *Consider two systems Σ and $\hat{\Sigma}$ as shown in Figure 4.3, where the outputs are specified by the values inside the brackets associated to each state, and secret states are marked in red. First note that one can easily verify that the smaller system $\hat{\Sigma}$ is δ-approximate initial-state opaque with $\delta = 0.1$. Next, we show that Σ is ε-approximate InitSOP simulated by $\hat{\Sigma}$, as in Definition 4.3.1, through the relation $R = \{(A, J), (B, K), (C, K), (D, K), (E, N), (F, M), (G, M), (I, M)\}$, where $\varepsilon = 0.1$. Condition 1 in Definition 4.3.1 can be easily checked since : a) for $E \in X_0 \cap X_S$, there exists $N \in \hat{X}_0 \cap \hat{X}_S$ such that $(E, N) \in R$; b) for $J \in \hat{X}_0 \setminus \hat{X}_S$, there exists $A \in X_0 \setminus X_S$ such that $(A, J) \in R$. Condition 2 is satisfied readily by seeing $\mathbf{d}(h(x), \hat{h}(\hat{x})) \leq 0.1$ holds for any $(x, \hat{x}) \in R$. One can also verify that condition 3 holds as well by checking*

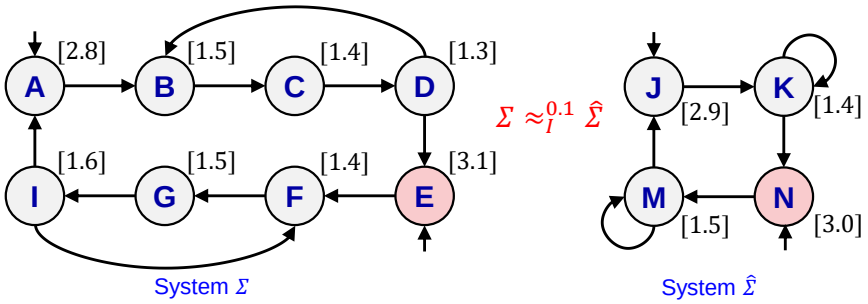


Figure 4.3: Example of ε-approximate initial-state opacity-preserving simulation relation.

conditions 3a) and 3b) for each pair of states in the relation R. For instance, consider the state pair $(C, K) \in R$, we have for $C \longrightarrow D$, there exists $K \longrightarrow K$, such that $(D, K) \in R$, and vice versa. Hence, R is an ε-InitSOP simulation relation from Σ to $\hat{\Sigma}$ as in Definition 4.3.1. Now, without applying any verification algorithm to Σ, by leveraging the results in Theorem 4.3.2, we can readily conclude that Σ is 0.3-approximate initial-state opaque, where $0.3 = \delta + 2\varepsilon$.

4.3.2 Approximate Current-State opacity-preserving Simulation Relation

Now, we provide a notion of approximate simulation relation for preserving current-state opacity.

Definition 4.3.6. *(Approximate Current-State opacity-preserving Simulation Relation) Consider two metric systems $\Sigma = (X, X_0, X_S, U, f, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h})$ with the same output sets $Y = \hat{Y}$ and metric $\mathbf{d}$. For $\varepsilon \in \mathbb{R}_{\geq 0}$, a relation $R \subseteq X \times \hat{X}$ iscalled an ε-approximate current-state opacity-preserving simulation relation (ε-CurSOP simulation relation) from Σ to $\hat{\Sigma}$ if*

1. $\forall x_0 \in X_0, \exists \hat{x}_0 \in \hat{X}_0 : (x_0, \hat{x}_0) \in R$;
2. $\forall (x, \hat{x}) \in R : \mathbf{d}(h(x), \hat{h}(\hat{x})) \leq \varepsilon$;
3. *For any $(x, \hat{x}) \in R$, we have*
   a) $\forall x \xrightarrow{u} x', \exists \hat{x} \xrightarrow{\hat{u}} \hat{x}' : (x', \hat{x}') \in R$;
   b) $\forall x \xrightarrow{u} x' \in X_S, \exists \hat{x} \xrightarrow{\hat{u}} \hat{x}' \in \hat{X}_S : (x', \hat{x}') \in R$;
   c) $\forall \hat{x} \xrightarrow{\hat{u}} \hat{x}', \exists x \xrightarrow{u} x' : (x', \hat{x}') \in R$;
   d) $\forall \hat{x} \xrightarrow{\hat{u}} \hat{x}' \in \hat{X} \setminus \hat{X}_S, \exists x \xrightarrow{u} x' \in X \setminus X_S : (x', \hat{x}') \in R$.

We say that Σ is ε-CurSOP simulated by $\hat{\Sigma}$, denoted by $\Sigma \preceq_C^{\varepsilon} \hat{\Sigma}$, if there exists an ε-CurSOP simulation relation R from Σ to $\hat{\Sigma}$.

The following theorem provides a sufficient condition for δ-approximate current-state opacity based on related systems as in Definition 4.3.6.

Theorem 4.3.7. Consider two metric systems $\Sigma = (X, X_0, X_S, U, f, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h})$ with the same output sets $Y = \hat{Y}$ and metric $\mathbf{d}$ and let $\varepsilon, \delta \in R_{\geq 0}$. If $\Sigma \preceq_C^{\varepsilon} \hat{\Sigma}$ and $\varepsilon \leq \frac{\delta}{2}$, then we have:

$$\begin{aligned} &\hat{\Sigma} \text{ is } (\delta - 2\varepsilon)\text{-approximate current-state opaque} \\ \Rightarrow &\Sigma \text{ is } \delta\text{-approximate current-state opaque.} \end{aligned}$$

Proof. Let us consider an arbitrary initial state $x_0 \in X_0$ and finite run $\mathbf{x} = x_0 x_1 \cdots x_n$ in Σ such that $x_n \in \Sigma$. We consider the following two cases: $n = 0$ and $n \neq 0$. If $n = 0$, we know that $x_0 \in \Sigma$. Since we assume that $\{x \in X_0 : (h(x_0), h(x)) \leq \delta\} \not\subseteq \Sigma$, we observe immediately that there exists $x_0' \in X_0 \setminus X_S$ such that $\mathbf{d}(h(x_0), h(x_0')) \leq \delta$. Then, we consider the case of $n \geq 1$. Since $\Sigma \preceq_C^{\varepsilon} \hat{\Sigma}$, by conditions 1, 2, 3a) and 3b) in Definition 4.3.6, there exist an initial state $\hat{x}_0 \in \hat{X}_0$ and a finite run $\hat{\mathbf{x}} = \hat{x}_0 \hat{x}_1 \cdots \hat{x}_n$ in $\hat{\Sigma}$ such that $\hat{x}_n \in \hat{X}_S$ and

$$\forall i \in \{0, 1, \ldots, n\} : \mathbf{d}(h(x_i), \hat{h}(x_i)) \leq \varepsilon. \tag{4.3.5}$$

Since $\hat{\Sigma}$ is $(\delta - 2\varepsilon)$-approximate current-state opaque, there exist an initial state $\hat{x}_0' \in \hat{X}_0$ and a finite run $\hat{\mathbf{x}}' = \hat{x}_0' \hat{x}_1' \cdots \hat{x}_n'$ such that $\hat{x}_n' \in \hat{X} \setminus \hat{X}_S$ and

$$\max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(\hat{h}(\hat{x}_i), \hat{h}(\hat{x}_i')) \leq \delta - 2\varepsilon. \tag{4.3.6}$$

Again, since $\Sigma \preceq_C^{\varepsilon} \hat{\Sigma}$, by conditions 1, 2, 3c) and 3d) in Definition 4.3.6, there exist an initial state $x_0' \in X_0$ and a finite run $\mathbf{x}' = x_0' x_1' \cdots x_n'$ in Σ such that $x_n' \in X \setminus X_S$ and

$$\forall i \in \{0, 1, \ldots, n\} : \mathbf{d}(h(x_i'), \hat{h}(\hat{x}_i')) \leq \varepsilon. \tag{4.3.7}$$

Combining equations (4.3.5), (4.3.6), (4.3.7), and using the triangle inequality, we have

$$\max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x_i')) \leq \delta. \tag{4.3.8}$$

Since $x_0 \in X_0$ and the finite run $\mathbf{x} = x_0 x_1 \cdots x_n$ are arbitrary, we conclude that Σ is δ-approximate current-state opaque. □

4.3.3 Approximate Infinite-Step opacity-preserving Simulation Relation

Finally, by combing ε-CurSOP simulation relation and ε-InitSOP simulation relation, we provide a notion of approximate simulation relation for preserving infinite-step opacity.

Definition 4.3.8. *(Approximate Infinite-Step opacity-preserving Simulation Relation) Consider two metric systems* $\Sigma = (X, X_0, X_S, U, f, Y, h)$ *and* $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h})$ *with the same output sets* $Y = \hat{Y}$ *and metric* $\mathbf{d}$*. For* $\varepsilon \in \mathbb{R}_{\geq 0}$*, a relation* $R \subseteq X \times \hat{X}$ *is called an* ε*-approximate infinite-step opacity-preserving simulation relation (*ε*-InfSOP simulation relation) from* Σ *to* $\hat{\Sigma}$ *if it is both an* ε*-CurSOP simulation relation from* Σ *to* $\hat{\Sigma}$ *and an* ε*-InitSOP simulation relation from* Σ *to* $\hat{\Sigma}$*, i.e.,*

1. *a)* $\forall x_0 \in X_0, \exists \hat{x}_0 \in \hat{X}_0 : (x_0, \hat{x}_0) \in R$*;*
   b) $\forall x_0 \in X_0 \cap X_S, \exists \hat{x}_0 \in \hat{X}_0 \cap \hat{X}_S : (x_0, \hat{x}_0) \in R$*;*
   c) $\forall \hat{x}_0 \in \hat{X}_0 \setminus \hat{X}_S, \exists x_0 \in X_0 \setminus X_S : (x_0, \hat{x}_0) \in R$*;*
2. $\forall (x, \hat{x}) \in R : \mathbf{d}(h(x), \hat{h}(\hat{x})) \leq \varepsilon$*;*
3. *For any* $(x, \hat{x}) \in R$*, we have*
   a) $\forall x \xrightarrow{u} x', \exists \hat{x} \xrightarrow{\hat{u}} \hat{x}' : (x', \hat{x}') \in R$*;*
   b) $\forall x \xrightarrow{u} x' \in X_S, \exists \hat{x} \xrightarrow{\hat{u}} \hat{x}' \in \hat{X}_S : (x', \hat{x}') \in R$*;*
   c) $\forall \hat{x} \xrightarrow{\hat{u}} \hat{x}', \exists x \xrightarrow{u} x' : (x', \hat{x}') \in R$*;*
   d) $\forall \hat{x} \xrightarrow{\hat{u}} \hat{x}' \in \hat{X} \setminus \hat{X}_S, \exists x \xrightarrow{u} x' \in X \setminus X_S : (x', \hat{x}') \in R$*.*

We say that Σ *is* ε*-InfSOP simulated by* $\hat{\Sigma}$*, denoted by* $\Sigma \preceq^{\varepsilon}_{IF} \hat{\Sigma}$*, if there exists an* ε*-InfSOP simulation relation* R *from* Σ *to* $\hat{\Sigma}$*.*

Similar to the cases of initial-state opacity and current-state opacity, we have the following theorem as a sufficient condition for δ-approximate infinite-step opacity based on related systems as in Definition 4.3.8.

Theorem 4.3.9. Consider two metric systems $\Sigma = (X, X_0, X_S, U, f, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h})$ with the same output sets $Y = \hat{Y}$ and metric $\mathbf{d}$ and let $\varepsilon, \delta \in R_{\geq 0}$. If $\Sigma \preceq^{\varepsilon}_{IF} \hat{\Sigma}$ and $\varepsilon \leq \frac{\delta}{2}$, then the following implication hold:

$$\begin{aligned} &\hat{\Sigma} \text{ is } (\delta - 2\varepsilon)\text{-approximate infinite-step opaque} \\ \Rightarrow &\Sigma \text{ is } \delta\text{-approximate infinite-step opaque.} \end{aligned}$$

Proof. Let us consider an arbitrary initial state $x_0 \in X_0$ and finite run $\mathbf{x} = x_0 x_1 \cdots x_n$ in Σ such that $x_k \in \Sigma$ for some $k = 0, \ldots, n$. We consider the following two cases:

If $k = 0$, then we have $x_0 \in \Sigma$. Since $\Sigma \preceq^{\varepsilon}_{IF} \hat{\Sigma}$ implies $\Sigma \preceq^{\varepsilon}_{I} \hat{\Sigma}$, by the proof of Theorem 4.3.2, we know that there exist an initial state $x'_0 \in X_0 \setminus X_S$ and a run $\mathbf{x}' = x'_0 x'_1 \cdots x'_n$ in Σ such that $\max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x'_i)) \leq \delta$.

If $k \geq 1$, then similar to the proof of Theorem 4.3.7, by conditions 1a), 2, 3a), 3b), 3c) and 3d) in Definition 4.3.8 and the fact the $\hat{\Sigma}$ is $(\delta - 2\varepsilon)$-approximate infinite-step opaque, there exist an initial state $x'_0 \in X_0$ and a finite run $\mathbf{x}' = x'_0 x'_1 \cdots x'_n$ such that $x'_k \in X \setminus X_S$ and $\max_{i \in \{0,1,\ldots,n\}} \mathbf{d}(h(x_i), h(x'_i)) \leq \delta$.

Since $x_0 \in X_0$, $\mathbf{x} = x_0x_1\cdots x_n$ and index k are arbitrary, we conclude that Σ is δ-approximate infinite-step opaque. □

4.4 Opacity of Discrete-Time Control Systems

In the previous section, we introduced notions of approximate opacity-preserving simulation relation and discussed their properties. This allows us to verify approximate opacity of infinite systems, e.g., continuous dynamical systems, based on their finite abstractions. In general, how to construct symbolic abstractions are system-dependent and not all systems admit symbolic models. In this section, we show that a class of discrete-time control systems do admit symbolic models for the purpose of verifying approximate opacity under certain stability assumptions.

4.4.1 Construction of Opacity-Preserving Finite Abstractions for Discrete-Time Control Systems

Now, we introduce a finite abstraction (a.k.a. symbolic system) for a discrete-time control system $\Sigma = (X, X_0, X_S, U, f, Y, h)$ as in Definition 2.3.1. To do so, from now on we assume that sets X, X_S and U are of the form of finite union of boxes. Assume that the output map h satisfies the following general Lipschitz assumption: $\|h(x)-h(x')\| \leq \alpha(\|x - x'\|)$, for all $x, x' \in X$, where $\alpha \in \mathcal{K}_\infty$. Consider a tuple $\mathsf{q} = (\eta, \mu, \theta)$ of parameters, where $0 < \eta \leq \min\{span(X_S), span(X \setminus X_S)\}$ is the state set quantization, $0 < \mu \leq span(U)$ is the input set quantization parameter, and θ is a design parameter. A finite abstraction of Σ is defined as

$$\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h}), \tag{4.4.1}$$

where $\hat{X} = \hat{X}_0 = [X]_\eta$, $\hat{X}_S = \left[X_S^\theta\right]_\eta$, $\hat{U} = [U]_\mu$, $\hat{Y} = \{h(\hat{x}) \mid \hat{x} \in \hat{X}\}$, where $\hat{h}(\hat{x}) = h(\hat{x})$, $\forall \hat{x} \in \hat{X}$, and

- $\hat{x}' \in \hat{f}(\hat{x}, \hat{u})$ if and only if $\|\hat{x}' - f(\hat{x}, \hat{u})\| \leq \eta$.

The following result shows that, for the class of δ-ISS control system as in Definition 2.4.1, under some condition over the quantization parameters η and μ, $\hat{\Sigma}$ and Σ are related under the approximate InitSOP simulation relation as in Definition 4.3.1.

Theorem 4.4.1 (**Initial-State Opacity-Preserving Finite Abstractions**). Consider a δ-ISS control system $\Sigma = (X, X_0, X_S, U, f, Y, h)$. For any desired precision $\varepsilon > 0$, let $\hat{\Sigma}$ be a finite abstraction of Σ as in (4.4.1) with a tuple $\mathsf{q} = (\eta, \mu, 0)$ of parameters satisfying

$$\beta\left(\alpha^{-1}(\varepsilon), 1\right) + \gamma(\mu) + \eta \leq \alpha^{-1}(\varepsilon), \tag{4.4.2}$$

then, we have $\Sigma \preceq_I^\varepsilon \hat{\Sigma} \preceq_I^\varepsilon \Sigma$.

Proof. We start by proving $\Sigma \preceq_I^{\varepsilon} \hat{\Sigma}$. Consider the relation $R \subseteq X \times \hat{X}$ defined by $(x, \hat{x}) \in R$ if and only if $\|x - \hat{x}\| \leq \alpha^{-1}(\varepsilon)$. Since $\eta \leq \mathit{span}(X_S)$, $X_S \subseteq \bigcup_{p \in [X_S]_\eta} \mathcal{B}_\eta(p)$, and by (4.4.2), $\forall x \in X_S$, $\exists \hat{x} \in \hat{X}_S$ such that:

$$\|x - \hat{x}\| \leq \eta \leq \alpha^{-1}(\varepsilon). \tag{4.4.3}$$

Hence, $(x, \hat{x}) \in R$ and condition 1a) in Definition 4.3.1 is satisfied. For every $\hat{x} \in \hat{X} \setminus \hat{X}_S$, by choosing $x = \hat{x}$ which is also inside set $X \setminus X_S$, one gets $(x, \hat{x}) \in R$ and, hence, condition 1b) in Definition 4.3.1 holds as well. Now consider any $(x, \hat{x}) \in R$. Condition 2 in Definition 4.3.1 is satisfied by the definition of R and the Lipschitz assumption:

$$\|h(x) - \hat{h}(\hat{x})\| = \|h(x) - h(\hat{x})\| \leq \alpha(\|x - \hat{x}\|) \leq \varepsilon.$$

Let us now show that condition 3 in Definition 4.3.1 holds.

Consider any $u \in U$. Choose an input $\hat{u} \in \hat{U}$ satisfying:

$$\|u - \hat{u}\| \leq \mu. \tag{4.4.4}$$

Note that the existence of such $\hat{u}$ is guaranteed by the inequality $\mu \leq \mathit{span}(U)$ which guarantees that $U \subseteq \bigcup_{p \in [U]_\mu} \mathcal{B}_\mu(p)$. Consider the transition $x' = f(x, u)$ in Σ. It follows from the δ-ISS assumption on Σ and (4.4.4) that the distance between x' and $f(\hat{x}, \hat{u})$ is bounded as:

$$\begin{aligned} \|x' - f(\hat{x}, \hat{u})\| \leq & \beta\left(\|x - \hat{x}\|, 1\right) + \gamma\left(\|u - \hat{u}\|\right) \\ \leq & \beta\left(\alpha^{-1}(\varepsilon), 1\right) + \gamma\left(\mu\right). \end{aligned} \tag{4.4.5}$$

Since $X \subseteq \bigcup_{p \in [X]_\eta} \mathcal{B}_\eta(p)$, there exists $\hat{x}' \in \hat{X}$ such that:

$$\|f(\hat{x}, \hat{u}) - \hat{x}'\| \leq \eta, \tag{4.4.6}$$

which, by the definition of $\hat{\Sigma}$, implies the existence of $\hat{x}' \in \hat{f}(\hat{x}, \hat{u})$ in $\hat{\Sigma}$. Using the inequalities (4.4.2), (4.4.5), (4.4.6), and triangle inequality, we obtain:

$$\begin{aligned} \|x' - \hat{x}'\| &\leq \|x' - f(\hat{x}, \hat{u}) + f(\hat{x}, \hat{u}) - \hat{x}'\| \\ &\leq \|x' - f(\hat{x}, \hat{u})\| + \|f(\hat{x}, \hat{u}) - \hat{x}'\| \\ &\leq \beta\left(\alpha^{-1}(\varepsilon), 1\right) + \gamma\left(\mu\right) + \eta \leq \alpha^{-1}(\varepsilon). \end{aligned}$$

Therefore, we conclude $(x', \hat{x}') \in R$ and condition 3a) in Definition 4.3.1 holds. Let us now show that condition 3b) in Definition 4.3.1 also holds.

Now consider any $(x, \hat{x}) \in R$ and any $\hat{u} \in \hat{U}$. Choose the input $u = \hat{u}$ and consider the unique $x' = f(x, u)$ in Σ. Using δ-ISS assumption for Σ, we bound the distance between x' and $f(\hat{x}, \hat{u})$ as:

$$\|x' - f(\hat{x}, \hat{u})\| \leq \beta\left(\|x - \hat{x}\|, 1\right) \leq \beta\left(\alpha^{-1}(\varepsilon), 1\right). \tag{4.4.7}$$

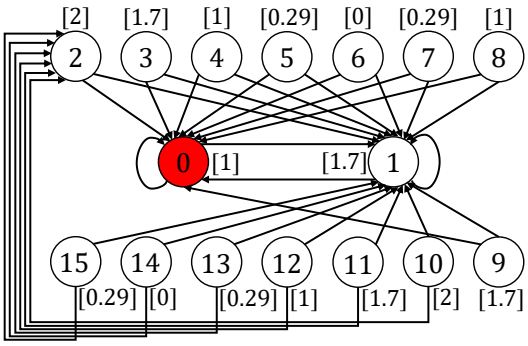


Figure 4.4: Symbolic model $\hat{\Sigma}$ associated with control systems Σ in (4.4.8) with $\eta = 0.1, \mu = 0.001$, and $\varepsilon = 0.9$.

Using the definition of $\hat{\Sigma}$, the inequalities (4.4.2), (4.4.7), and the triangle inequality, we obtain:

$$\begin{aligned}\|x' - \hat{x}'\| \leq & \|x' - f(\hat{x}, \hat{u}) + f(\hat{x}, \hat{u}) - \hat{x}'\| \\ \leq & \|x' - f(\hat{x}, \hat{u})\| + \|f(\hat{x}, \hat{u}) - \hat{x}'\| \\ \leq & \beta\left(\alpha^{-1}(\varepsilon), 1\right) + \eta \leq \alpha^{-1}(\varepsilon).\end{aligned}$$

Therefore, we conclude that $(x', \hat{x}') \in R$ and condition 3b) in Definition 4.3.1 holds.

In a similar way, one can prove that $\hat{\Sigma} \preceq_I^{\varepsilon} \Sigma$. □

Remark 4.4.2. *Note that there always exist quantization parameters* q *such that inequality* (4.4.2) *holds as long as* $\beta\left(\alpha^{-1}(\varepsilon), 1\right) < \alpha^{-1}(\varepsilon)$. *By assuming that the discrete-time control system* Σ *is a sampled-data version of an original continuous-time one with the sampling time* τ, *one can ensure the latter inequality by choosing the sampling time large enough given that* $\beta(r, 1) = \hat{\beta}(r, \tau) < r$ *for some* $\mathcal{KL}$ *function* $\hat{\beta}$ *establishing the incremental stability of the original continuous-time system. For example, for the function in* (2.4.3), *one has* $\beta(r, 1) = \|A\| r = \|\mathsf{e}^{\hat{A}\tau}\| r$, *where* $\hat{A}$ *is the state matrix of the original continuous-time linear control system.*

The following example illustrates how to use Theorem 4.4.1 to verify approximate opacity for an infinite system based on its finite abstraction.

Example 4.4.3. *Let us consider the following simple system*

$$\Sigma : \begin{cases} \xi(k+1 & = 0.1\xi(k) + \upsilon(k), \\ \zeta(k) & = \sin(2.5\pi\xi(k)) + 1, \end{cases} \tag{4.4.8}$$

where $X = [0, 1.6[, X_S = [0, 0.1[$ *and* $U = \{0.001\}$. *This system is clearly* δ*-ISS and according to Equation* (2.4.3), *we have* $\beta(r, k) = 0.1^k r$ *and* $\gamma(r) = \sum_{m=0}^{\infty} 0.1^m r$. *Also, function* h *satisfies the Lipschitz condition with* $\alpha(r) = 2.5\pi r$. *By Equation* (4.4.2), *the parameters* $\mathsf{q} = (\eta, \mu, 0)$ *and the abstract precision* ε *should satisfy* $\frac{0.04}{\pi}\varepsilon + \frac{10}{9}\mu + \eta \leq$

$\frac{0.4}{\pi}\varepsilon$. *Let us consider desired abstract precision* $\varepsilon = 0.9$ *and quantization parameters* $\mathsf{q} = (\eta, \mu, 0) = (0.1, 0.001, 0)$ *satisfying the inequality. Then we obtain symbolic system* $\hat{\Sigma}$ *shown in Figure 4.4, and by Theorem 4.4.1, we have* $\Sigma \preceq_I^{0.9} \hat{\Sigma} \preceq_I^{0.9} \Sigma$. *Essentially, we discretize the state space of* $[0, 1.6[$ *into* 16 *discrete states based on parameter* η. *One can easily check that* $\hat{\Sigma}$ *is* 0*-approximate initial-state opaque since for any run from secret initial state* 0*, there exists a run from non-secret state* 8 *such that their outputs are exactly the same. Therefore, by Theorem 4.3.2, we can conclude that* Σ *is* 1.8*-approximate initial-state opaque.*

The next theorem provides similar results as in Theorem 4.4.1 but by leveraging δ-ISS Lyapunov functions. To show the next result, we will make the following supplementary assumption on the δ-ISS Lyapunov functions as in Definition 2.4.3. We assume that there exists a function $\hat{\gamma} \in \mathcal{K}_\infty$ such that

$$\forall x, x', x'' \in X, \quad V(x, x') - V(x', x'') \leq \hat{\gamma}(\|x - x''\|). \tag{4.4.9}$$

Inequality (4.4.9) is not restrictive at all provided we are interested in the dynamics of the control system on a compact subset of the state set X; see the discussion in [53].

Theorem 4.4.4. *Let* $\Sigma = (X, X_0, X_S, U, f, Y, h)$ *admit a* δ*-ISS Lyapunov function* V *satisfying* (4.4.9)*. For any desired precision* $\varepsilon > 0$*, let* $\hat{\Sigma}$ *be a finite abstraction of* Σ *as in* (4.4.1) *with a tuple* $\mathsf{q} = (\eta, \mu, 0)$ *of quantization parameters satisfying*

$$\overline{\alpha}(\eta) \leq \underline{\alpha}(\alpha^{-1}(\varepsilon)), \tag{4.4.10}$$

$$\max\{\kappa(\underline{\alpha}(\alpha^{-1}(\varepsilon))), \lambda(\mu)\} + \hat{\gamma}(\eta) \leq \underline{\alpha}(\alpha^{-1}(\varepsilon)), \tag{4.4.11}$$

then we have $\Sigma \preceq_I^{\varepsilon} \hat{\Sigma} \preceq_I^{\varepsilon} \Sigma$.

Proof. We start by proving $\Sigma \preceq_I^{\varepsilon} \hat{\Sigma}$. Consider the relation $R \subseteq X \times \hat{X}$ defined by $(x, \hat{x}) \in R$ if and only if $V(x, \hat{x}) \leq \underline{\alpha}(\alpha^{-1}(\varepsilon))$. Since $\eta \leq \mathit{span}(X_S)$ and $X_S \subseteq \bigcup_{p \in [X_S]_\eta} \mathcal{B}_\eta(p)$, for every $x \in X_S$ there always exists $\hat{x} \in \hat{X}_S$ such that $\|x - \hat{x}\| \leq \eta$. Then

$$V(x, \hat{x}) \leq \overline{\alpha}(\|x - \hat{x}\|) \leq \overline{\alpha}(\eta) \leq \underline{\alpha}(\alpha^{-1}(\varepsilon))$$

because of (4.4.10) and $\overline{\alpha}$ being a $\mathcal{K}_\infty$ function. Hence, $(x, \hat{x}) \in R$ and condition 1a) in Definition 4.3.1 is satisfied. For every $\hat{x} \in \hat{X} \setminus \hat{X}_S$, by choosing $x = \hat{x}$ which is also inside set $X \setminus X_S$, one gets trivially $(x, \hat{x}) \in R$ and, hence, condition 1b) in Definition 4.3.1 holds as well. Now consider any $(x, \hat{x}) \in R$. Condition 2 in Definition 4.3.1 is satisfied by the definition of R and the Lipschitz assumption on map h:

$$\begin{aligned} \|h(x) - \hat{h}(\hat{x})\| = \|h(x) - h(\hat{x})\| &\leq \alpha(\|x - \hat{x}\|) \\ &\leq \alpha(\underline{\alpha}^{-1}(V(x, \hat{x})) \leq \varepsilon. \end{aligned}$$

Let us now show that condition 3 in Definition 4.3.1 holds.

Consider any $u \in U$. Choose an input $\hat{u} \in \hat{U}$ satisfying:

$$\|u - \hat{u}\| \leq \mu. \tag{4.4.12}$$

Note that the existence of such $\hat{u}$ is guaranteed by the inequality $\mu \leq span(U)$ which guarantees that $U \subseteq \bigcup_{p\in[U]_\mu} \mathcal{B}_\mu(p)$. Consider the unique transition $x \xrightarrow{u} x' = f(x,u)$ in Σ. Given δ-ISS Lyapunov function V for Σ, inequality (2.4.4), and (4.4.12), one obtains:

$$\begin{aligned} V(x', f(\hat{x},\hat{u})) &\leq \max\{\kappa\left(V(x,\hat{x})\right), \lambda\left(\|u-\hat{u}\|\right)\} \\ &\leq \max\{\kappa\left(\underline{\alpha}(\alpha^{-1}(\varepsilon))\right), \lambda\left(\mu\right)\}. \end{aligned} \tag{4.4.13}$$

Since $X \subseteq \bigcup_{p\in[X]_\eta} \mathcal{B}_\eta(p)$, there exists $\hat{x}' \in \hat{X}$ such that:

$$\|f(\hat{x},\hat{u}) - \hat{x}'\| \leq \eta, \tag{4.4.14}$$

which, by the definition of $\hat{\Sigma}$, implies the existence of $\hat{x}' \in \hat{f}(\hat{x},\hat{u})$ in $\hat{\Sigma}$. Using the inequalities (4.4.9), (4.4.11), (4.4.13), and (4.4.14), we obtain:

$$\begin{aligned} V(x',\hat{x}') &\leq V(x', f(\hat{x},\hat{u})) + \hat{\gamma}(\|f(\hat{x},\hat{u}) - \hat{x}'\|) \\ &\leq \max\{\kappa\left(\underline{\alpha}(\alpha^{-1}(\varepsilon))\right), \lambda\left(\mu\right)\} + \hat{\gamma}\left(\eta\right) \\ &\leq \underline{\alpha}(\alpha^{-1}(\varepsilon)). \end{aligned}$$

Therefore, we conclude $(x',\hat{x}') \in R$ and condition 3a) in Definition 4.3.1 holds. Let us now show that condition 3b) in Definition 4.3.1 also holds.

Now consider any $(x,\hat{x}) \in R$. Consider any $\hat{u} \in \hat{U}$. Choose the input $u = \hat{u}$ and consider the unique $x' = f(x,u)$ in Σ. Given δ-ISS Lyapunov function V for Σ and inequality (2.4.4), one gets:

$$V(x', f(\hat{x},\hat{u})) \leq \kappa\left(V(x,\hat{x})\right) \leq \kappa\left(\underline{\alpha}(\alpha^{-1}(\varepsilon))\right). \tag{4.4.15}$$

Using the definition of $\hat{\Sigma}$, the inequalities (4.4.9), (4.4.11), and (4.4.15), we obtain:

$$\begin{aligned} V(x',\hat{x}') \leq& V(x', f(\hat{x},\hat{u})) + \hat{\gamma}(\|f(\hat{x},\hat{u}) - \hat{x}'\|) \\ \leq& \kappa\left(\underline{\alpha}(\alpha^{-1}(\varepsilon))\right) + \hat{\gamma}(\eta) \leq \underline{\alpha}(\alpha^{-1}(\varepsilon)). \end{aligned}$$

Therefore, we conclude that $(x',\hat{x}') \in R$ and condition 3b) in Definition 4.3.1 holds.

In a similar way, one can prove that $\hat{\Sigma} \preceq^{\varepsilon}_{I} \Sigma$. □

Remark 4.4.5. *One can readily verify that there always exists a choice of quantization parameter* $\mathsf{q} = (\eta, \mu, 0)$ *such that inequalities* (4.4.10) *and* (4.4.11) *hold simultaneously. Although the result in Theorem 4.4.4 seems more general than that of Theorem 4.4.1 in terms of the existence of quantization parameter* q*, the symbolic model* $\hat{\Sigma}$*, computed by using the quantization parameters* q *provided in Theorem 4.4.1 whenever existing, is likely to have fewer states than the one computed by using the quantization parameters provided in Theorem 4.4.4 due to the conservative nature of* δ*-ISS Lyapunov functions.*

Remark 4.4.6. *The notions of approximate opacity are, in general, hard to check for a concrete system since there is no systematic way in the literature to check opacity for systems with infinite state set so far. On the other hand, existing tool* `DESUMA`[1] *and algorithms [209], [162],[222, Sec. IV] in DESs literature can be leveraged to check* exact *opacity for systems with finite state sets. For the verification of approximate opacity of the constructed finite abstractions, one can readily resort to Section 4.2 for an effective verification approach that was developed for the notion of* approximate opacity *for finite systems.*

The next theorems illustrate another main results of this section showing that, under similar conditions over the quantization parameters η and μ, $\hat{\Sigma}$ and Σ are related under an approximate current-state opacity-preserving simulation relation.

Theorem 4.4.7 (**Current-State Opacity-Preserving Finite Abstractions**)**.** Let $\Sigma = (X, X_0, X_S, U, f, Y, h)$ be a δ-ISS control system. For any desired precision $\varepsilon > 0$, let $\hat{\Sigma}$ be a finite abstraction of Σ as in (4.4.1) with a tuple $\mathsf{q} = (\eta, \mu, \theta)$ of parameters satisfying

$$\beta\left(\alpha^{-1}(\varepsilon), 1\right) + \gamma(\mu) + \eta \leq \alpha^{-1}(\varepsilon),$$
$$\alpha^{-1}(\varepsilon) \leq \theta,$$

then we have $\Sigma \preceq^{\varepsilon}_{C} \hat{\Sigma}$.

Proof. Consider the relation $R \subseteq X \times \hat{X}$ defined by $(x, \hat{x}) \in R$ if and only if $\|x - \hat{x}\| \leq \alpha^{-1}(\varepsilon)$. Note that conditions 1, 2, 3a) and 3c) of ε-CurSOP simulation relation in Definition 4.3.6 are similar to that of ε-InitSOP simulation relation, therefore the proof of them are similar to that in Theorem 4.4.1 and is omitted here. Here, we show that conditions 3b) and 3d) in Definition 4.3.6 hold.

Let us consider an arbitrary transition $x' = f(x, u)$ with $x' \in X_S$ in Σ. Similar to the proof of condition 3a), we can show the existence of a transition $\hat{x}' \in \hat{f}(\hat{x}, \hat{u})$ in $\hat{\Sigma}$ where $(x', \hat{x}') \in R$ holds, where the input $\hat{u} \in \hat{U}$ satisfies: $\|u - \hat{u}\| \leq \mu$. By the construction of the secret set in the symbolic system, one has $\hat{X}_S = [X^{\theta}_S]_{\eta}$ with $\theta \geq \alpha^{-1}(\varepsilon)$ and $0 < \eta \leq \min\{span(X_S), span(X \setminus X_S)\}$. Therefore, since $(x', \hat{x}') \in R$ which implies $\|x' - \hat{x}'\| \leq \alpha^{-1}(\varepsilon)$, we obtain that $\hat{x}' \in \hat{X}_S$. Thus, we conclude that condition 3b) in Definition 4.3.6 holds. In a similar way, we can show that condition 3d) in Definition 4.3.6 holds as well which completes the proof. □

Theorem 4.4.8. *Let $\Sigma = (X, X_0, X_S, U, f, Y, h)$ admits a δ-ISS Lyapunov function V satisfying* (4.4.9)*. For any desired precision $\varepsilon > 0$, let $\hat{\Sigma}$ be a finite abstraction of Σ as in* (4.4.1) *with a tuple $\mathsf{q} = (\eta, \mu, \theta)$ of parameters satisfying*

$$\overline{\alpha}(\eta) \leq \underline{\alpha}(\alpha^{-1}(\varepsilon)),$$

[1] Available at URL http://www.eecs.umich.edu/umdes/toolboxes.html.

$$\begin{aligned}\max\{\kappa(\underline{\alpha}(\alpha^{-1}(\varepsilon))),\lambda(\mu)\}+\hat{\gamma}(\eta) \leq& \underline{\alpha}(\alpha^{-1}(\varepsilon)),\\ \alpha^{-1}(\varepsilon) \leq& \theta,\end{aligned}$$

then we have $\Sigma \preceq^{\varepsilon}_{C} \hat{\Sigma}$.

Proof. The proof is similar to that of Theorem 4.4.4 and Theorem 4.4.7 and is omitted here. □

Since we show $\Sigma \preceq^{\varepsilon}_{I} \hat{\Sigma}$ and $\Sigma \preceq^{\varepsilon}_{C} \hat{\Sigma}$ under the *same* relation in Theorems 4.4.1 and 4.4.7 (resp. Theorems 4.4.4 and 4.4.8), by the definition of approximate infinite-state opacity-preserving simulation relation, we consequently get the following results where the proofs are omitted.

Theorem 4.4.9. *Let* $\Sigma = (X, X_0, X_S, U, f, Y, h)$ *be a* δ*-ISS control system. For any desired precision* $\varepsilon > 0$*, let* $\hat{\Sigma}$ *be a finite abstraction of* Σ *as in* (4.4.1) *with a tuple* $\mathsf{q} = (\eta, \mu, \theta)$ *of parameters satisfying*

$$\begin{aligned}\beta\left(\alpha^{-1}(\varepsilon), 1\right)+\gamma(\mu)+\eta \leq& \alpha^{-1}(\varepsilon),\\ \alpha^{-1}(\varepsilon) \leq& \theta,\end{aligned}$$

then we have $\Sigma \preceq^{\varepsilon}_{IF} \hat{\Sigma}$.

Theorem 4.4.10. *Let* $\Sigma = (X, X_0, X_S, U, f, Y, h)$ *admits a* δ*-ISS Lyapunov function* V *satisfying* (4.4.9)*. For any desired precision* $\varepsilon > 0$*, let* $\hat{\Sigma}$ *be a finite abstraction of* Σ *as in* (4.4.1) *with a tuple* $\mathsf{q} = (\eta, \mu, \theta)$ *of parameters satisfying*

$$\begin{aligned}\overline{\alpha}(\eta) \leq& \underline{\alpha}(\alpha^{-1}(\varepsilon)),\\ \max\{\kappa(\underline{\alpha}(\alpha^{-1}(\varepsilon))),\lambda(\mu)\}+\hat{\gamma}(\eta) \leq& \underline{\alpha}(\alpha^{-1}(\varepsilon)),\\ \alpha^{-1}(\varepsilon) \leq& \theta,\end{aligned}$$

then we have $\Sigma \preceq^{\varepsilon}_{IF} \hat{\Sigma}$.

4.5 Opacity of Discrete-Time Stochastic Control Systems

In the previous section, we provided approximate simulation relations that preserve approximate opacity for the class of (non-stochastic) discrete-time control systems. We also discussed how to construct finite abstractions that approximately simulate a class of discrete-time control systems in terms of opacity preservation. The results bridge the gap between the opacity analysis of finite discrete systems and continuous control systems. However, in real-world applications, a small probability of violation of the opacity could be tolerable. Hence, instead of simply asking if a system is opaque or non-opaque, it is more applicable to evaluate the possibility of being not opaque for stochastic settings. We address the problem of abstraction-based opacity verification of discrete-time stochastic control systems (dt-SCS) in this section.

4.5.1 Opacity-Preserving Stochastic Simulation Functions

In this subsection, we introduce a notion of initial-state opacity-preserving stochastic simulation functions for dt-SCS. The stochastic simulation function will play an important role in analyzing opacity for dt-SCS. First, we provide the definition of initial-state opacity-preserving stochastic simulation functions.

Definition 4.5.1. (Initial-state opacity-preserving stochastic simulation function) Let $\Sigma = (X, X_0, X_S, U, \varsigma, f, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \varsigma, \hat{f}, \hat{Y}, \hat{h})$ be two dt-SCS with the same output sets $Y = \hat{Y}$. A function $V : X \times \hat{X} \to \mathbb{R}_{\geq 0}$ is called an initial-state opacity-preserving stochastic simulation function (InitSOP-SSF) from $\hat{\Sigma}$ to Σ, if there exist constants $\psi \geq 0$, $\omega \geq 0$, a function $\alpha \in \mathcal{K}_\infty$, and a function $\kappa \in \mathcal{K}$ which satisfies $\kappa(s) \geq \hat{\kappa} s$, $\forall s \in \mathbb{R}_{\geq 0}$, where $0 < \hat{\kappa} < 1$, such that

1 a) $\forall x_0 \in X_0 \cap X_S$, $\exists \hat{x}_0 \in \hat{X}_0 \cap \hat{X}_S$: $V(x_0, \hat{x}_0) \leq \omega$;
b) $\forall \hat{x}_0 \in \hat{X}_0 \setminus \hat{X}_S$, $\exists x_0 \in X_0 \setminus X_S$: $V(x_0, \hat{x}_0) \leq \omega$;

2 $\forall x \in X, \forall \hat{x} \in \hat{X}$, $\alpha(\|h(x) - \hat{h}(\hat{x})\|) \leq V(x, \hat{x})$;

3 $\forall x \in X, \forall \hat{x} \in \hat{X}$, the following conditions hold:
a) $\forall u, \quad \exists \hat{u}, \quad \text{s.t.} \quad \mathbb{E}\Big[V(f(x,u,\varsigma), \hat{f}(\hat{x},\hat{u},\varsigma)) \,\Big|\, x, \hat{x}, u, \hat{u}\Big] - V(x,\hat{x}) \leq -\kappa(V(x,\hat{x})) + \psi$;
b) $\forall \hat{u}, \quad \exists u, \quad \text{s.t.} \quad \mathbb{E}\Big[V(f(x,u,\varsigma), \hat{f}(\hat{x},\hat{u},\varsigma)) \,\Big|\, x, \hat{x}, u, \hat{u}\Big] - V(x,\hat{x}) \leq -\kappa(V(x,\hat{x})) + \psi$.

Now, before stating the main theorem of this section, we provide the following technical proposition which is inspired by Theorem 3.3 in [97]. This proposition shows us the usefulness of the InitSOP-SSF in the sense that it can be employed to show indistinguishability of output trajectories of two dt-SCS in a probabilistic setting.

Proposition 4.5.2. *Let $\Sigma = (X, X_0, X_S, U, \varsigma, f, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \varsigma, \hat{f}, \hat{Y}, \hat{h})$ be two dt-SCS with the same output sets $Y = \hat{Y}$. Suppose V is an InitSOP-SSF from $\hat{\Sigma}$ to Σ. Then, for any $a \in X_0 \cap X_S$ in Σ, there exists $\hat{a} \in \hat{X}_0 \cap \hat{X}_S$ in $\hat{\Sigma}$ (respectively, for any $\hat{a} \in \hat{X}_0 \setminus \hat{X}_S$ in $\hat{\Sigma}$, there exists $a \in X_0 \setminus X_S$ in Σ) so that for any $\hat{\nu} \in \hat{\mathcal{U}}$ in $\hat{\Sigma}$, there exists $\nu \in \mathcal{U}$ in Σ and vice versa such that the following inequality holds*

$$\mathbb{P}\left\{\sup_{0\leq k\leq n} \|\zeta_{a\nu}(k) - \hat{\zeta}_{\hat{a}\hat{\nu}}(k)\| \leq \lambda \,|\, [a;\hat{a}]\right\} \geq 1 - \bar{\varepsilon}_\lambda,$$
$$\bar{\varepsilon}_\lambda := \begin{cases} 1 - (1 - \frac{\omega}{\alpha(\lambda)})(1 - \frac{\psi}{\alpha(\lambda)})^n & \text{if } \alpha(\lambda) \geq \frac{\psi}{\hat{\kappa}}, \\ (\frac{\omega}{\alpha(\lambda)})(1-\hat{\kappa})^n + (\frac{\psi}{\hat{\kappa}\alpha(\lambda)})(1 - (1-\hat{\kappa})^n) & \text{if } \alpha(\lambda) < \frac{\psi}{\hat{\kappa}}, \end{cases} \tag{4.5.1}$$

for any $\lambda > 0$.

Proof. It can be readily seen that by conditions 2 and 3 in Definition 4.5.1, the InitSOP-SSF is a stochastic simulation function (SSF) (as defined in [97, Definition 3.2]) both

from $\hat{\Sigma}$ to Σ and from Σ to $\hat{\Sigma}$. Since by condition 1 in Definition 4.5.1, $V(a,\hat{a}) \leq \omega$, the rest of the proof is concluded by applying Theorem 3.3 in [97]. □

This proposition will be used for the proof of the following main theorem, where we show preservation of approximate initial-state opacity across related systems as in Definition 4.5.1. The next lemmas will be used to prove the main result.

Lemma 4.5.3. *Suppose for two dt-SCS Σ and $\hat{\Sigma}$, the output trajectories $\zeta_{a\nu}$ and $\hat{\zeta}_{\hat{a}\hat{\nu}}$ satisfy the inequality*

$$\sup_{0\leq k\leq n} \|\zeta_{a\nu}(k) - \hat{\zeta}_{\hat{a}\hat{\nu}}(k)\| \leq \lambda,$$

for some time bound n and $\lambda > 0$. Then we have:

$$\hat{\zeta}_{\hat{a}\hat{\nu}} \in \underline{E}_\lambda \Longrightarrow \zeta_{a\nu} \in E; \zeta_{a\nu} \in E \Longrightarrow \hat{\zeta}_{\hat{a}\hat{\nu}} \in \bar{E}^\lambda,$$

over time interval $[0,n]$, for any measurable set $E \subseteq B_Y$ and the modified sets $\underline{E}_\lambda$ and $\bar{E}^\lambda$ as defined in (3.3.2) *and* (3.3.3)*.*

Proof. As can be seen from the definition of $\underline{E}_\lambda$ and $\bar{E}^\lambda$ in (3.3.2) and (3.3.3), given any set of output sequences $E \subseteq B_Y$, $\underline{E}_\lambda$ and $\bar{E}^\lambda$ are the λ-deflated version and λ-inflated version of set E, respectively. Since we have

$$\sup_{0\leq k\leq n} \|\zeta_{a\nu}(k) - \hat{\zeta}_{\hat{a}\hat{\nu}}(k)\| \leq \lambda,$$

then it can be readily seen that according to the structure of $\underline{E}_\lambda$ and $\bar{E}^\lambda$, $\hat{\zeta}_{\hat{a}\hat{\nu}} \in \underline{E}_\lambda$ guarantees $\zeta_{a\nu} \in E$. Similarly, $\zeta_{a\nu} \in E$ inplies $\hat{\zeta}_{\hat{a}\hat{\nu}} \in \bar{E}^\lambda$ as well. □

This lemma essentially provides us the relation between the property satisfactions of two dt-SCS, given that the output trajectories of these two dt-SCS are close to each other. Based on this lemma, the following lemma presents another technical result of this section.

Lemma 4.5.4. *Suppose Σ and $\hat{\Sigma}$ are two dt-SCS for which inequality* (4.5.1) *holds with initial states a and $\hat{a}$, input sequences ν and $\hat{\nu}$, a constant pair $(\lambda, \bar{\varepsilon}_\lambda)$ and any time bound n. The following inequality holds for any set $E \subseteq B_Y$ and the modified sets $\underline{E}_\lambda$ and $\bar{E}^\lambda$ as defined in* (3.3.2) *and* (3.3.3)*:*

$$\mathbb{P}(\hat{\zeta}_{\hat{a}\hat{\nu}} \in \underline{E}_\lambda) - \bar{\varepsilon}_\lambda \leq \mathbb{P}(\zeta_{a\nu} \in E) \leq \mathbb{P}(\hat{\zeta}_{\hat{a}\hat{\nu}} \in \bar{E}^\lambda) + \bar{\varepsilon}_\lambda, \tag{4.5.2}$$

where the satisfaction is over time interval $\{0,\ldots,n\}$.

Proof. Let us consider the events:

$$\mathcal{E}_1 := \{\hat{\zeta}_{\hat{a}\hat{\nu}} \in \underline{E}_\lambda\}, \mathcal{E}_2 := \{\zeta_{a\nu} \in E\}, \mathcal{E}_3 := \{\sup_{0\leq k\leq n} \|\zeta_{a\nu}(k) - \hat{\zeta}_{\hat{a}\hat{\nu}}(k)\| \leq \lambda\}.$$

Since we have from Lemma 4.5.3, $\mathcal{E}_1 \cap \mathcal{E}_3 \implies \mathcal{E}_2$, thus, $\mathbb{P}(\bar{\mathcal{E}}_2) \leq \mathbb{P}(\bar{\mathcal{E}}_1 \cup \bar{\mathcal{E}}_3) \leq \mathbb{P}(\bar{\mathcal{E}}_1) + \mathbb{P}(\bar{\mathcal{E}}_3)$, where $\bar{\mathcal{E}}_1$, $\bar{\mathcal{E}}_2$ and $\bar{\mathcal{E}}_3$ are the complements of $\mathcal{E}_1$, $\mathcal{E}_2$ and $\mathcal{E}_3$, respectively. As we have by (4.5.1), $\mathbb{P}(\bar{\mathcal{E}}_3) \leq \bar{\varepsilon}_\lambda$, now we readily get:

$$\begin{aligned}\mathbb{P}(\bar{\mathcal{E}}_2) \leq \mathbb{P}(\bar{\mathcal{E}}_1) + \bar{\varepsilon}_\lambda &\implies 1 - \mathbb{P}(\mathcal{E}_2) \leq 1 - \mathbb{P}(\mathcal{E}_1) + \bar{\varepsilon}_\lambda \\ &\implies \mathbb{P}(\mathcal{E}_1) \leq \mathbb{P}(\mathcal{E}_2) + \bar{\varepsilon}_\lambda,\end{aligned}$$

which gives us $\mathbb{P}(\hat{\zeta}_{\hat{a}\hat{\nu}} \in \underline{E}_\lambda) - \bar{\varepsilon}_\lambda \leq \mathbb{P}(\zeta_{a\nu} \in E)$. The proof of $\mathbb{P}(\zeta_{a\nu} \in E) \leq \mathbb{P}(\hat{\zeta}_{\hat{a}\hat{\nu}} \in \bar{E}^\lambda) + \bar{\varepsilon}_\lambda$ is similar and is omitted here due to lack of space. □

Now, we present the main result of this section on the preservation of opacity across related dt-SCS systems.

Theorem 4.5.5. Let $\Sigma = (X, X_0, X_S, U, \varsigma, f, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \varsigma, \hat{f}, \hat{Y}, \hat{h})$ be two dt-SCS with the same output sets $Y = \hat{Y}$. Consider constants $\varepsilon \in \mathbb{R}_{\geq 0}$ and $\lambda \in \mathbb{R}_{>0}$. Assume V is an InitSOP-SSF from $\hat{\Sigma}$ to Σ as in Definition 4.5.1 with the corresponding constants ψ, ω, $\hat{\kappa}$ and $\mathcal{K}_\infty$ function α. Then the following implication holds:

$$\begin{aligned}&\hat{\Sigma} \text{ is } \varepsilon\text{-approximate initial-state opaque} \\ \Rightarrow &\Sigma \text{ is } (2\lambda, \varepsilon + 2\bar{\varepsilon}_\lambda)\text{-approximate initial-state opaque,}\end{aligned} \tag{4.5.3}$$

where $\bar{\varepsilon}_\lambda \in \mathbb{R}_{\geq 0}$ is computed as in (4.5.1).

Proof. Consider an arbitrary secret initial state $x_0 \in X_0 \cap X_S$, input sequence $\nu = \{u_1, u_2, \ldots, u_n\}$ and the corresponding state run $\xi_{x_0\nu} = (x_0, x_1, \ldots, x_n)$ in Σ. Since V is an InitSOP-SSF from $\hat{\Sigma}$ to Σ, by conditions 1a), 2 and 3a) in Definition 4.5.1, there exist a secret initial state $\hat{x}_0 \in \hat{X}_0 \cap \hat{X}_S$, input sequence $\hat{\nu} = \{\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_n\}$ and state run $\hat{\xi}_{\hat{x}_0\hat{\nu}} = (\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_n)$ in $\hat{\Sigma}$ such that $V(x_0, \hat{x}_0) \leq \omega$, and $\forall i \in \{0, 1, \ldots, n\}$:

$$\begin{aligned}\alpha(\|h(x_i) - \hat{h}(\hat{x}_i)\|) &\leq V(x_i, \hat{x}_i), \\ \mathbb{E}\Big[V(f(x_i, u_i, \varsigma_i)), \hat{f}(\hat{x}_i, \hat{u}_i, \varsigma_i)) \,\Big|\, x_i, \hat{x}_i, u_i, \hat{u}_i\Big] - V(x_i, \hat{x}_i) &\leq -\kappa(V(x_i, \hat{x}_i)) + \psi.\end{aligned}$$

By applying Proposition 4.5.2, for the given λ, we have:

$$\mathbb{P}\left\{\max_{0 \leq i \leq n} \|\zeta_{x_0\nu}(i) - \zeta_{\hat{x}_0\hat{\nu}}(i)\| \leq \lambda \,|\, [x_0; \hat{x}_0]\right\} \geq 1 - \bar{\varepsilon}_\lambda,$$

where $\bar{\varepsilon}_\lambda$ is computed using inequality (4.5.1). By applying (4.5.2) in Lemma 4.5.4, we get for any set $E \subseteq B_Y$ and the modified sets $\underline{E}_\lambda$ and $\bar{E}^\lambda$:

$$\mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in \underline{E}_\lambda) - \mathbb{P}(\zeta_{x_0\nu} \in E) \leq \bar{\varepsilon}_\lambda, \tag{4.5.4}$$

$$\mathbb{P}(\zeta_{x_0\nu} \in E) - \mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in \bar{E}^\lambda) \leq \bar{\varepsilon}_\lambda. \tag{4.5.5}$$

Since $\hat{\Sigma}$ is ε-approximate initial-state opaque, by Definition 3.3.1, there exist a non-secret initial state $\hat{x}'_0 \in \hat{X}_0 \setminus \hat{X}_S$, input sequence $\hat{\nu}' = \{\hat{u}'_1, \hat{u}'_2, \ldots, \hat{u}'_n\}$ and state run $\xi_{\hat{x}'_0\hat{\nu}'} = (\hat{x}'_0, \hat{x}'_1, \ldots, \hat{x}'_n)$ in $\hat{\Sigma}$, such that $\|\mathbb{P}(\hat{\zeta}_{\hat{x}'_0\hat{\nu}'} \in E) - \mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in E)\| \leq \varepsilon$ holds for any set E, so we have:

$$\mathbb{P}(\hat{\zeta}_{\hat{x}'_0\hat{\nu}'} \in \underline{E}_\lambda) - \mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in \underline{E}_\lambda) \leq \varepsilon, \tag{4.5.6}$$

$$\mathbb{P}(\hat{\zeta}_{\hat{x}_0\hat{\nu}} \in \bar{E}^\lambda) - \mathbb{P}(\hat{\zeta}_{\hat{x}'_0\hat{\nu}'} \in \bar{E}^\lambda) \leq \varepsilon. \tag{4.5.7}$$

Again, since V is an InitSOP-SSF from $\hat{\Sigma}$ to Σ, by conditions 1b), 2 and 3b) in Definition 4.5.1, Proposition 4.5.2 and (4.5.2) in Lemma 4.5.4, there exist an initial state $x'_0 \in X_0 \setminus X_S$, input sequence $\nu' = \{u'_1, u'_2, \ldots, u'_n\}$ and the corresponding state run $\xi_{x'_0\nu'} = (x'_0, x'_1, \ldots, x'_n)$ in Σ such that

$$\mathbb{P}(\zeta_{x'_0\nu'} \in \underline{E}_{2\lambda}) - \mathbb{P}(\hat{\zeta}_{\hat{x}'_0\hat{\nu}'} \in \underline{E}_\lambda) \leq \bar{\varepsilon}_\lambda, \tag{4.5.8}$$

$$\mathbb{P}(\hat{\zeta}_{\hat{x}'_0\hat{\nu}'} \in \bar{E}^\lambda) - \mathbb{P}(\zeta_{x'_0\nu'} \in \bar{E}^{2\lambda}) \leq \bar{\varepsilon}_\lambda. \tag{4.5.9}$$

Hence, by combining inequalities (4.5.4), (4.5.6), (4.5.8), we have the following result

$$\mathbb{P}(\zeta_{x'_0\nu'} \in \underline{E}_{2\lambda}) - \mathbb{P}(\zeta_{x_0\nu} \in E) \leq \varepsilon + 2\bar{\varepsilon}_\lambda. \tag{4.5.10}$$

Additionally, combining inequalities (4.5.5), (4.5.7), (4.5.9), we get

$$\mathbb{P}(\zeta_{x_0\nu} \in E) - \mathbb{P}(\zeta_{x'_0\nu'} \in \bar{E}^{2\lambda}) \leq \varepsilon + 2\bar{\varepsilon}_\lambda. \tag{4.5.11}$$

Since $x_0 \in X_0 \cap X_S$ and input sequence ν in Σ are arbitrary, we conclude that Σ is $(2\lambda, \varepsilon + 2\bar{\varepsilon}_\lambda)$-approximate initial-state opaque. □

Remark 4.5.6. *This theorem provides a sufficient condition for approximate initial-state opacity based on the relation between two stochastic systems. It bridges the gap between the verification of opacity and abstraction-based techniques for stochastic systems. By constructing an abstraction of system Σ, which appears as system $\hat{\Sigma}$ in the theorem, and leveraging the simulation relation between them, one can efficiently verify opacity of the complex system Σ. The abstraction is contructed as a finite Markov decision process (MDP) in the following Subsection 4.5.2.1. In addition, as mentioned in Remark 3.3.7, ε-approximate initial-state opacity for the MDP can be verified easily using existing computation algorithms for total variation distance.*

4.5.2 Construction of Opacity-Preserving Finite Abstractions for Discrete-Time Stochastic Control Systems

In this section, we show how to analyze approximate opacity for the class of dt-SCS based on their finite abstractions (finite MDPs). First, we provide the construction of finite abstractions of the concrete systems.

4.5.2.1 Finite Abstractions of Discrete-Time Stochastic Control Systems

Given a dt-SCS $\Sigma = (X, X_0, X_S, U, \varsigma, f, Y, h)$, we construct a finite MDP as its finite abstraction, represented by the tuple $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \varsigma, \hat{f}, \hat{Y}, \hat{h})$. The construction of finite MDPs follows a similar procedure as in [97, Algorithm 1] with some modifications to incorporate the role of secret sets. First, for the given state set X and input set U, we select finite partitions of them as $X = \cup_i \mathsf{X}_i, U = \cup_i \mathsf{U}_i$, and single representative points $\bar{x}_i \in \mathsf{X}_i$, $\bar{u}_i \in \mathsf{U}_i$ as abstract states and inputs. Then, we define the state and input sets of $\hat{\Sigma}$ as $\hat{X} = \{\bar{x}_i, i = 1, \ldots, n_x\}$, and $\hat{U} = \{\bar{u}_i, i = 1, \ldots, n_u\}$, which simply consist of the selected representative points. The transition function $\hat{f} : \hat{X} \times \hat{U} \times V_\varsigma \to \hat{X}$ is defined as

$$\hat{f}(\hat{x}, \hat{u}, \varsigma) = \Pi_x(f(\hat{x}, \hat{u}, \varsigma)), \tag{4.5.12}$$

where $\Pi_x : X \to \hat{X}$ represents the map that assigns to any $x \in X$, the representative point $\hat{x} \in \hat{X}$ of the corresponding partition set containing x. The output set $\hat{Y}$ is the image of $\hat{X}$ under h, with $\hat{h} : \hat{X} \to \hat{Y}$ being the same as h except for having a restricted domain $\hat{X}$. Similarly, we use $\Pi_u : U \to \hat{U}$ to denote the map that assigns to any $u \in U$, the representative input point $\hat{u} \in \hat{U}$ of the corresponding partition set containing u.

Remark 4.5.7. *Note that we have not defined the map $\Pi_x : X \to \hat{X}$ yet. For example, one can choose center points (if applicable) of the partitions as representative points or apply other specialized mapping rule to it. In this work, we enforce two conditions as the rule of choosing representative points for initial states and secret states as follows:*

1 *If $X_0 \cap \mathsf{X}_i \neq \emptyset$, we constrain the representative point of X_i to be chosen as $\bar{x}_i \in X_0 \cap \mathsf{X}_i$;*

2 *If $X_S \cap \mathsf{X}_i \neq \emptyset$, we constrain the representative point of X_i to be chosen as $\bar{x}_i \in X_S \cap \mathsf{X}_i$.*

By the above conditions, one can observe that the initial state set $\hat{X}_0$ and secret state set $\hat{X}_S$ satisfy $\hat{X}_0 \subseteq X_0$ and $\hat{X}_S \subseteq X_S$.

Remark 4.5.8. *In this section, it is assumed that the abstraction maps Π_x and Π_u satisfy the inequalities*

$$\|\Pi_x(x) - x\| \leq \mu_x, \forall x \in X, \|\Pi_u(u) - u\| \leq \mu_u, \forall u \in U, \tag{4.5.13}$$

where μ_x and μ_u are the state *and* input *discretization parameter defined as*

$$\mu_x := \sup\{\|x - x'\|,\ x, x' \in \mathsf{X}_i,\ i = 1, 2, \ldots, n_x\}, \tag{4.5.14}$$
$$\mu_u := \sup\{\|u - u'\|,\ u, u' \in \mathsf{U}_i,\ i = 1, 2, \ldots, n_u\}. \tag{4.5.15}$$

Next, we construct the InitSOP-SSF for a class of nonlinear stochastic systems.

4.5.2.2 Establishing InitSOP-SSF for a Class of Nonlinear Stochastic Systems

In this subsection, we focus on a general class of nonlinear stochastic systems Σ. We provide an InitSOP-SSF candidate for the concrete systems Σ and their finite MDPs as constructed in the previous subsection. The existence of such an InitSOP-SSF enables us to verify opacity of a continuous-space stochastic system by leveraging its finite abstraction. The establishment of InitSOP-SSF is under the following two assumptions. First, we assume that the output map h satisfies the following general Lipschitz assumption: there exists an $\tilde{\alpha} \in \mathcal{K}_\infty$ such that $\|h(x) - h(x')\| \leq \tilde{\alpha}(\|x - x'\|)$ for all $x, x' \in X$. Second, we assume that the concrete system is δ-ISS as in the following definition. Note that this definition is a variant of Definition 2.4.3 tailored to dt-SCS.

Definition 4.5.9. *A dt-SCS Σ is* incrementally input-to-state stable *(δ-ISS) if there exists function $V : X \times X \to \mathbb{R}_{\geq 0}$ such that $\forall x, x' \in X$, $\forall u, u' \in U$ the following two inequalities hold:*

$$\underline{\alpha}(\|x - x'\|) \leq V(x, x') \leq \overline{\alpha}(\|x - x'\|), \tag{4.5.16}$$

$$\mathbb{E}\Big[V(f(x, u, \varsigma), f(x', u', \varsigma))\big|x, x', u, u'\Big] - V(x, x') \leq -\bar{\kappa}(V(x, x')) + \bar{\rho}(\|u - u'\|), \tag{4.5.17}$$

for some $\underline{\alpha}, \overline{\alpha} \in \mathcal{K}_\infty$, $\bar{\kappa} \in \mathcal{K}$, and $\bar{\rho} \in \mathcal{K}_\infty \cup \{0\}$.

Now, we provide the main theorem in this subsection. We show that by adding a mild condition, the function V described in Definition 4.5.9 is indeed an InitSOP-SSF from the finite abstraction $\hat{\Sigma}$ (as constructed in Subsection 4.5.2.1) to the concrete system Σ.

Theorem 4.5.10. Consider a δ-ISS dt-SCS Σ that admits a function V as in Definition 4.5.9. Let $\hat{\Sigma}$ be its *finite* MDP constructed as in Subsection 4.5.2.1. Suppose there exists a constant $0 < \hat{\kappa} < 1$ such that the function $\bar{\kappa} \in \mathcal{K}$ satisfies $\bar{\kappa}(s) \geq \hat{\kappa} s$, $\forall s \in \mathbb{R}_{\geq 0}$. Assume that there exists a function $\gamma \in \mathcal{K}_\infty$ such that V satisfies

$$V(x, x') - V(x, x'') \leq \gamma(\|x' - x''\|), \forall x, x', x'' \in X. \tag{4.5.18}$$

Then V is an InitSOP-SSF from $\hat{\Sigma}$ to Σ.

Proof. We start by proving condition 1 in Definition 4.5.1. For every initial state $x_0 \in X_0 \cap X_S$ in Σ, there always exists a representative point $\hat{x}_0 = \Pi_x(x_0)$ in $\hat{\Sigma}$ which is inside the set $\hat{X}_0 \cap \hat{X}_S$ by the construction of $\hat{X}_0$ and $\hat{X}_S$, and $\|\hat{x}_0 - x_0\| \leq \mu_x$ holds by (4.5.13). Hence, we have $V(x_0, \hat{x}_0) \leq \overline{\alpha}(\|x_0 - \hat{x}_0\|)$ by (4.5.16), and condition 1a) in Definition 4.5.1 is satisfied with $\omega = \overline{\alpha}(\mu_x)$. For every $\hat{x}_0 \in \hat{X}_0 \setminus \hat{X}_S$, by choosing $x_0 = \hat{x}_0$ which is also inside $X_0 \setminus X_S$, we get $V(x_0, \hat{x}_0) = 0 \leq \omega$. Hence, condition 1b) in Definition 4.5.1 holds as well.

Let us show condition 2 in Definition 4.5.1 holds. Since Σ is incrementally input-to-state stable and using (4.5.16), and given the Lipschitz assumption on h, $\forall x \in X$ and $\forall \hat{x} \in \hat{X}$, one gets

$$\|h(x) - \hat{h}(\hat{x})\| \leq \tilde{\alpha}(\|x - \hat{x}\|) \leq \tilde{\alpha} \circ \underline{\alpha}^{-1}(V(x, \hat{x})),$$

which results in

$$\alpha(\|h(x) - \hat{h}(\hat{x})\|) \leq V(x, \hat{x}),$$

$\forall x \in X$ and $\forall \hat{x} \in \hat{X}$, where $\alpha(s) := (\tilde{\alpha} \circ \underline{\alpha}^{-1})^{-1}(s)$, $\forall s \in \mathbb{R}_{\geq 0}$. Hence, condition 2 in Definition 4.5.1 is satisfied. Let us now show that condition 3 in Definition 4.5.1 holds as well. Now, $\forall x \in X, \forall \hat{x} \in \hat{X}$, $\forall u \in U$ and $\forall \hat{u} \in \hat{U}$, by taking the conditional expectation from (4.5.18), we have

$$\begin{aligned}&\mathbb{E}\Big[V(f(x,u,\varsigma), \hat{f}(\hat{x},\hat{u},\varsigma)) \,\Big|\, x,\hat{x},u,\hat{u}\Big] - \mathbb{E}\Big[V(f(x,u,\varsigma), f(\hat{x},\hat{u},\varsigma)) \,\Big|\, x,\hat{x},u,\hat{u}\Big] \\ &\leq \mathbb{E}\Big[\gamma(\|\hat{f}(\hat{x},\hat{u},\varsigma) - f(\hat{x},\hat{u},\varsigma)\|) \,\Big|\, x,\hat{x},u,\hat{u}\Big].\end{aligned}$$

Employing (4.5.17), one gets

$$\mathbb{E}\Big[V(f(x,u,\varsigma), f(\hat{x},\hat{u},\varsigma)) \,\Big|\, x,\hat{x},u,\hat{u}\Big] \leq V(x,\hat{x}) - \bar{\kappa}(V(x,\hat{x})) + \bar{\rho}(\|u - \hat{u}\|).$$

Since $\hat{f}(\hat{x},\hat{u},\varsigma) = \Pi_x(f(\hat{x},\hat{u},\varsigma))$, by using (4.5.13), we get

$$\mathbb{E}\Big[\gamma(\|\hat{f}(\hat{x},\hat{u},\varsigma) - f(\hat{x},\hat{u},\varsigma)\|) \,\Big|\, x,\hat{x},u,\hat{u}\Big] \leq \gamma(\mu_x).$$

Now, consider any $u \in U$. By choosing the representative input $\hat{u} = \Pi_u(u)$, which satisfies $\|u - \hat{u}\| \leq \mu_u$, we obtain

$$\mathbb{E}\Big[V(f(x,u,\varsigma), \hat{f}(\hat{x},\hat{u},\varsigma)) \,\Big|\, x,\hat{x},u,\hat{u}\Big] - V(x,\hat{x}) \leq -\bar{\kappa}(V(x,\hat{x})) + \bar{\rho}(\mu_u) + \gamma(\mu_x).$$

Hence, condition 3a) in Definition 4.5.1 holds with $\psi = \bar{\rho}(\mu_u) + \gamma(\mu_x)$. Similarly, $\forall x \in X, \forall \hat{x} \in \hat{X}$, and $\forall \hat{u} \in \hat{U}$, by choosing $u = \hat{u}$, we have

$$\begin{aligned}&\mathbb{E}\Big[V(f(x,u,\varsigma), \hat{f}(\hat{x},\hat{u},\varsigma)) \,\Big|\, x,\hat{x},u,\hat{u}\Big] - V(x,\hat{x}) \\ &\quad\leq -\bar{\kappa}(V(x,\hat{x})) + \gamma(\mu_x) \leq -\bar{\kappa}(V(x,\hat{x})) + \psi.\end{aligned}$$

Therefore, condition 3b) in Definition 4.5.1 holds as well, and we conclude that V is an InitSOP-SSF from $\hat{\Sigma}$ to Σ. □

4.6 Discussion and Future Work

In this chapter, we discussed an abstraction-based framework for the verification of approximate opacity for both non-stochastic and stochastic control systems. Verification algorithms and approximate simulation relations that preserve approximate opacity were provided tailored to our new concept of approximate opacity. We also discussed how to construct finite abstractions that approximately simulate classes of discrete-time (stochastic) control systems in terms of opacity preservation. Our results bridge the gap between the opacity analysis of finite discrete systems and continuous control systems.

In the following, we further discuss some ongoing research topics and open problems.

Verification of General Notion of Opacity for CPS Existing works for opacity verification of general CPS mainly focus on particular types of opacity such as initial-state opacity or infinite-step one. For finite systems, the general notion of α-opacity as defined in Definition 3.5.1 can be verified using the observer-like structures when the security properties can be realized by ω-automata. However, for general CPS with infinite states, how to verify the general notion of α-opacity still needs developments. In particular, for the abstraction-based approach, one needs to identify suitable relation that preserves α-opacity. For the barrier-based approach, appropriate conditions for barrier certificates of α-opacity also need to be identified.

Opacity Verification for Larger Classes of CPS In the context of analyzing stochastic systems, the results provided in Section 4.5 made some initial steps towards abstraction-based opacity verification. It is meaningful to further extend our framework to cover more notions of opacity, e.g., K-step opacity, current-state opacity and infinite-step opacity. Efficient verification algorithms need to be developed to facilitate abstraction-based verification frameworks for stochastic control systems. Moreover, the aforementioned abstraction-based approaches for opacity verification of general CPS crucially depends on incremental ISS assumption. However, this assumption is rather restrictive for many practical systems. How to relax the stability assumption so that the verification techniques can be applied to more general classes of CPS is an interesting and important future direction. Also, it will be useful to develop opacity verification techniques, either using abstraction-based techniques, for more complex classes of CPS with time-delays or uncertainties. Also, in the problem formulation of opacity, the attacker is assumed to be able to access partial information-flow of the plant. However, for networked control systems, the information transmission between controllers and plants in the feedback loops may also be released to the intruder. There are some very recent works on the verification of opacity for networked control systems using finite-state models; see, e.g., [210, 206, 223, 108, 205]. However, existing works on formal verification of networked control system mainly focus on the mission requirements [217, 65, 144, 21] and to the best of our knowledge, there is no result on formal verification of opacity for general networked CPS.

Abstraction-Based Synthesis of Opacity for CPS The notions of opacity-preserving alternating simulation relations (ASR) proposed in [70] made the first step towards abstraction-based opacity synthesis for CPS. However, it has many limitations that need to be addressed in the future. First, the results in [70] are developed for particular types of state-based opacity. Similar to the verification problem, we also need to extend the results, particularly the underlying simulation relations, to the general case of α-opacity. Second, the opacity-preserving ASR belongs to the category of exact simulation. This condition, in general, is too strong for general CPS with continuous state-space. It is likely that there does not exist a finite symbolic model simulating the concrete system exactly. One possible direction to address this issue is to enforce approximate opacity rather than the exact version. To this end, one needs to consider the approximate ASR [147, 218] rather than the exact ASR. Third, existing results only support state-feedback controllers, i.e., the controller knows the current-state of the system precisely. As we discussed, an opacity-enforcing controller is observation-based in general. To address this issue, a possible solution is to use the output-feedback refinement relation (OFRR) [157, 85] instead of the ASR. How to suitably generalize the OFRR to preserve opacity is still an open problem. Finally, although opacity-preserving relations have been identified, there is no abstraction algorithm available so far for building finite abstractions based on the concrete systems with continuous-space dynamics that satisfy those relations. When the concrete system is δ-ISS, the abstraction can be done analogous to the case of verification. The major open problem is how to build opacity-preserving finite abstractions for the purpose of control without the stability assumption.

5 A Deductive Approach for Opacity Verification via Barrier Certificates

5.1 Introduction

The results discussed in Chapter 4 provides a systematic framework via abstraction-based techniques to deal with opacity verification for complex CPS. However, this methodology may suffer from scalability issues since it requires discretization of the state and input sets of the original system. In fact, current techniques on the construction of finite abstractions can scale up to a few variables but run out of time or memory when confronted with larger models. Motivated by this limitation, this chapter provides an alternative discretization-free approach for the formal verification of approximate opacity for CPS via a notion of barrier certificates.

5.1.1 Related Literature

Barrier certificates have shown to be a promising tool for the analysis of safety problems [149, 6, 7, 197, 76] and recently extended to deal with more general temporal logic specifications [76, 109, 9]. In particular, the seminal work in [148] introduced for the first time a notion of *barrier certificates* as a tool for safety verification of a class of hybrid systems. Later, the authors extended this work in [150] where safety and reachability are studied as a dual pair, and verification approaches for safety and reachability were proposed by searching for such barrier certificates using optimization techniques. However, there are very few works on verification of security properties using barrier certificates.

A recent attempt to analyze privacy of CPS using barrier certificates is made in [2]. A new notion of current-state opacity was considered there based on the belief space of the intruder. The privacy verification problem is cast into checking a safety property of the intruder's belief dynamics using barrier certificates. However, this framework is again limited to systems modeled by partially-observable Markov decision processes (POMDPs) with finite state sets.

5.1.2 Contributions

In this chapter, we develop, for the first time, a discretization-free approach for the formal verification of approximate opacity based on notions of barrier certificates. First, we introduce the so-called *augmented system* constructed by taking the product of a system with itself. Then, two new notions of so-called *augmented control barrier*

certificates (ACBC) are defined for the augmented systems. The first type of ACBC guarantees a *safety* property of the augmented system in the sense that there is no trajectory originating from a given initial region reaching a given unsafe set. Along with this, the initial and unsafe regions are designed in a specific form capturing the secret and initial sets of the original system. In this way, the existence of an ACBC provides us a sufficient condition ensuring that the original system is approximate initial-state (or infinite-step) opaque. Note that the failure in finding such an ACBC does not mean the system is not opaque. Therefore, we further present another type of ACBC which proves a *reachability* property of the augmented system. This type of ACBC can be utilized for showing that the original system starting from the initial set will eventually reach the unsafe region. This type of ACBC, on the other hand, provides a sufficient condition showing that the original system is lacking approximate initial-state (or infinite-step) opacity. Apart from the analysis of approximate initial-state and infinite-step opacity, we further investigate relationships between different notions of state-based opacity, i.e., initial-state, current-state, and K-step opacity, and study conditions under which one property may imply another one. Finally, we present a way to compute polynomial ACBC by means of sum-of-squares (SOS) programming, where the conditions required for the ACBC are reformulated as SOS constraints.

5.2 Augmented Control Sytems

Consider a dt-CS $\Sigma = (X, X_0, X_S, U, f, Y, h)$ as in Definition 2.3.1. We define the associated augmented system by

$$\Sigma \times \Sigma = (X \times X, X_0 \times X_0, X_S \times X_S, U \times U, f \times f, Y \times Y, h \times h), \tag{5.2.1}$$

which can be seen as the product of a dt-CS Σ and itself. For later use, we denote by $(x, \hat{x}) \in X \times X$ a pair of states in $\Sigma \times \Sigma$ and by $(\mathbf{x}_{x_0,\nu}, \mathbf{x}_{\hat{x}_0,\hat{\nu}})$ the state trajectory of $\Sigma \times \Sigma$ starting from $(x_0, \hat{x}_0)$ under input run $(\nu, \hat{\nu})$. We use $\mathcal{R} = X \times X$ to denote the augmented state space.

5.3 Augmented Control Barrier Certificates

In the sequel, we propose a technique that is sound in verifying approximate opacity for discrete-time control systems. Our approach is based on finding two types of barrier certificates as defined next.

Here, we first define a notion of barrier certificates that is constructed over the augmented system $\Sigma \times \Sigma$ and ensures a safety property for $\Sigma \times \Sigma$.

Proposition 5.3.1. *Consider a dt-CS Σ as in Definition 2.3.1, the associated augmented system $\Sigma \times \Sigma$, and sets $\mathcal{R}_0, \mathcal{R}_u \subseteq \mathcal{R}$. Suppose there exists a function $B : X \times X \to \mathbb{R}$ and constants $\underline{\epsilon}, \overline{\epsilon} \in \mathbb{R}$ with $\overline{\epsilon} > \underline{\epsilon}$ such that*

$$\forall (x, \hat{x}) \in \mathcal{R}_0, \qquad B(x, \hat{x}) \leq \underline{\epsilon}, \tag{5.3.1}$$

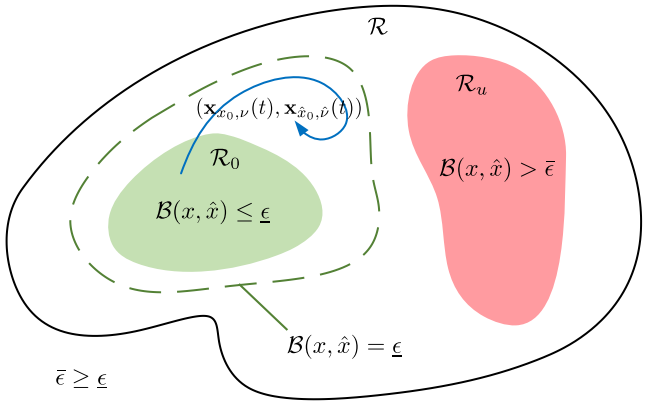


Figure 5.1: Barrier certificate ensuring safety of the augmented system, which implies opacity of the original system.

$$\forall (x,\hat{x}) \in \mathcal{R}_u, \qquad B(x,\hat{x}) \geq \overline{\epsilon}, \tag{5.3.2}$$

$$\forall (x,\hat{x}) \in \mathcal{R}, \forall u \in U, \exists \hat{u} \in U,$$
$$B(f(x,u), f(\hat{x},\hat{u})) - B(x,\hat{x}) \leq 0. \tag{5.3.3}$$

Then, for any initial condition $(x_0,\hat{x}_0) \in \mathcal{R}_0$ and for any input run ν, there exists an input run $\hat{\nu}$ such that $(\mathbf{x}_{x_0,\nu}(t), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(t)) \cap \mathcal{R}_u = \emptyset$, $\forall t \in \mathbb{N}$.

Proof. This proposition is proved by contradiction. Let us consider any state trajectory $(\mathbf{x}_{x_0,\nu}, \mathbf{x}_{\hat{x}_0,\hat{\nu}})$ of $\Sigma \times \Sigma$ that starts from an initial condition $(x_0,\hat{x}_0) \in \mathcal{R}_0$, under input sequences ν and $\hat{\nu}$. Suppose $\hat{\nu}$ is computed such that the inequality in (5.3.3) holds. Assume the state run reaches a state in $\mathcal{R}_u$, i.e., $(\mathbf{x}_{x_0,\nu}(t), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(t)) \in \mathcal{R}_u$ for some $t \in \mathbb{N}$. From (5.3.1) and (5.3.2), we have $B(x_0,\hat{x}_0) \leq \underline{\epsilon}$ and $B(\mathbf{x}_{x_0,\nu}(t), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(t)) \geq \overline{\epsilon}$. By using (5.3.3), one has $\overline{\epsilon} \leq B(\mathbf{x}_{x_0,\nu}(t), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(t)) \leq B(x_0,\hat{x}_0) \leq \underline{\epsilon}$, which contradicts $\overline{\epsilon} > \underline{\epsilon}$. Therefore, for any state trajectory of $\Sigma \times \Sigma$ starting from any initial condition in $\mathcal{R}_0$ under any input run ν, $(\mathbf{x}_{x_0,\nu}(t), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(t)) \cap \mathcal{R}_u = \emptyset$ always holds under the extracted control policy $\hat{\nu}$, which completes the proof. □

If $B(x,\hat{x})$ satisfies the conditions in Proposition 5.3.1, then it is called an *augmented control barrier certificate* (ACBC) for $\Sigma \times \Sigma$. The interpretation of a barrier certificate ensuring safety property of the augmented system is depicted in Figure 5.1.

Next, we further introduce another type of barrier certificates which ensures, on the other hand, the reachability property of the augmented system. Note that these two types of barrier certificates will be later used in reversed directions for the verification of opacity or lack of opacity of control systems.

Proposition 5.3.2. *Consider a dt-CS Σ as in Definition 2.3.1, the associated augmented system $\Sigma \times \Sigma$, and sets $\mathcal{R}_0, \mathcal{R}_u \subseteq \mathcal{R}$. Suppose $X \subset \mathbb{R}^n$ is a bounded set and there exists a continuous function $V : X \times X \to \mathbb{R}$ such that*

$$\forall (x,\hat{x}) \in \mathcal{R}_0, \qquad V(x,\hat{x}) \leq 0, \tag{5.3.4}$$

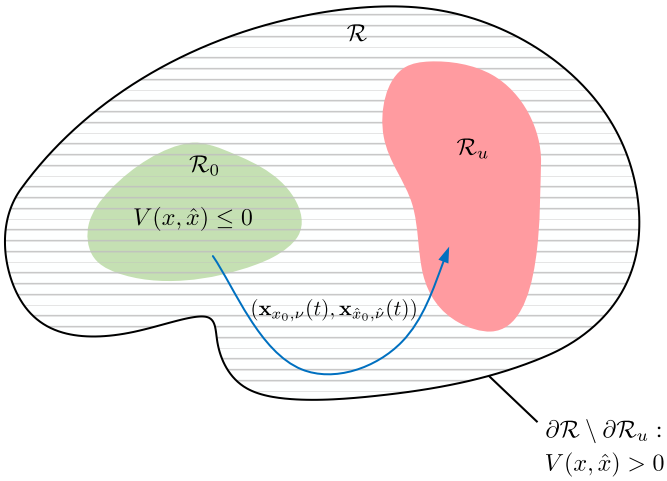


Figure 5.2: Barrier certificate ensuring reachability of the augmented system, which implies lack of opacity of the original system.

$$\forall (x,\hat{x}) \in \partial\mathcal{R} \setminus \partial\mathcal{R}_u, \qquad V(x,\hat{x}) > 0, \tag{5.3.5}$$

$$\forall (x,\hat{x}) \in \overline{(\mathcal{R} \setminus \mathcal{R}_u)}, \exists u \in U, \forall \hat{u} \in U, \\ V(f(x,u), f(\hat{x},\hat{u})) - V(x,\hat{x}) < 0. \tag{5.3.6}$$

Then, for any initial condition $(x_0,\hat{x}_0) \in \mathcal{R}_0$*, there exists an input run* ν *such that* $(\mathbf{x}_{x_0,\nu}(T), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(T)) \in \mathcal{R}_u$ *for any* $\hat{\nu}$*, for some* $T \geq 0$*, and* $(\mathbf{x}_{x_0,\nu}(t), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(t)) \in \mathcal{R}$*,* $\forall t \in [0;T]$.

Proof. Consider an initial state $(x_0,\hat{x}_0) \in \mathcal{R}_0$. One has $V(x_0,\hat{x}_0) \leq 0$ by (5.3.4). Consider an input run ν such that (5.3.6) is satisfied for the state runs $\mathbf{x}_{x_0,\nu}(t), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(t)$ of Σ, where $\hat{\nu}$ is an arbitrary input run. First note that the continuous function $V(x,\hat{x})$ is bounded below on the compact set $\overline{(\mathcal{R} \setminus \mathcal{R}_u)}$. From (5.3.6), $V(x,\hat{x})$ is strictly decreasing along the trajectory $(\mathbf{x}_{x_0,\nu}, \mathbf{x}_{\hat{x}_0,\hat{\nu}})$ in region $\overline{(\mathcal{R} \setminus \mathcal{R}_u)}$. It follows that $(\mathbf{x}_{x_0,\nu}, \mathbf{x}_{\hat{x}_0,\hat{\nu}})$ must leave $\overline{(\mathcal{R} \setminus \mathcal{R}_u)}$ in finite time. Now, assume $(\mathbf{x}_{x_0,\nu}, \mathbf{x}_{\hat{x}_0,\hat{\nu}})$ leaves $\overline{(\mathcal{R} \setminus \mathcal{R}_u)}$ without entering region $\mathcal{R}_u$ first. Consider the first time instant $t = T$ when $(\mathbf{x}_{x_0,\nu}(t), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(t))$ is leaving $\overline{(\mathcal{R} \setminus \mathcal{R}_u)}$, i.e., $(\mathbf{x}_{x_0,\nu}(t),\ \mathbf{x}_{\hat{x}_0,\hat{\nu}}(t)) \in \overline{(\mathcal{R} \setminus \mathcal{R}_u)}$ for all $t \in [0,T]$, and $(\mathbf{x}_{x_0,\nu}(T+\epsilon), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(T+\epsilon)) \notin \mathcal{R}$ for any $\epsilon > 0$. By (5.3.6) and $V(x_0,\hat{x}_0) \leq 0$, we have $V(\mathbf{x}_{x_0,\nu}(T), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(T)) \leq 0$ which contradicts (5.3.5). Therefore, we conclude that for any run starting from $\mathcal{R}_0$ under ν, there must exist $T \geq 0$ such that $(\mathbf{x}_{x_0,\nu}(T), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(T)) \in \mathcal{R}_u$ for any $\hat{\nu}$, and $(\mathbf{x}_{x_0,\nu}(t), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(t)) \in \mathcal{R}$, $\forall t \in [0,T]$, which completes the proof. □

A function $V(x,\hat{x})$ satisfying the conditions in Proposition 5.3.2 is also called an ACBC for $\Sigma \times \Sigma$. The idea of using barrier functions to prove reachability was first described in [150]. The interpretation of a barrier certificate ensuring reachability property of an augmented system is illustrated in Figure 5.2.

In the following section, we will describe how to use the above defined barrier certificates for the verification of different notions of approximate opacity for control systems.

5.4 Formal Verification of Opacity using Barrier Certificates

5.4.1 Verifying Approximate Initial-State Opacity

Here, we show how one can leverage the ACBC defined in the previous subsection to verify approximate initial-state opacity for a dt-CS Σ. To this purpose, we define the sets of initial conditions $\mathcal{R}_0$ and unsafe states $\mathcal{R}_u$ as:

$$\mathcal{R}_0 = \{(x, \hat{x}) \in (X_0 \cap X_S) \times (X_0 \setminus X_S) \mid \|h(x) - h(\hat{x})\| \leq \delta\}, \tag{5.4.1}$$
$$\mathcal{R}_u = \{(x, \hat{x}) \in X \times X \mid \|h(x) - h(\hat{x})\| > \delta\}, \tag{5.4.2}$$

where $\delta \in \mathbb{R}_{\geq 0}$ captures the measurement precision of the intruder as introduced in Definition 3.3.1. The following theorem provides us a sufficient condition in verifying approximate initial-state opacity of discrete-time control systems.

Theorem 5.4.1. *Consider a dt-CS Σ as in Definition 2.3.1. Suppose there exists a function $B : X \times X \to \mathbb{R}$ satisfying* (5.3.1)-(5.3.3) *in Proposition 5.3.1 with sets $\mathcal{R}_0, \mathcal{R}_u$ given in* (5.4.1)-(5.4.2)*. Then, system Σ is δ-approximate initial-state opaque.*

Proof. Consider an arbitrary secret initial state $x_0 \in X_0 \cap X_S$, any input run ν, and the corresponding state run $\mathbf{x}_{x_0,\nu}$ in Σ. First note that by (3.3.1), $\{x \in X_0 \mid \|h(x) - h(x_0)\| \leq \delta\} \not\subseteq X_S$. It follows that there exists an initial state $\hat{x}_0 \in X_0 \setminus X_S$ such that $\|h(\hat{x}_0) - h(x_0)\| \leq \delta$. Consider the pair of initial states $(x_0, \hat{x}_0)$. It can be readily seen that $(x_0, \hat{x}_0) \in \mathcal{R}_0$ as in (5.4.1). Now, given the existence of an ACBC as in Proposition 5.3.1, there exists a control policy $\hat{\nu}$ such that (5.3.3) is satisfied. By using Proposition 5.3.1, under $\hat{\nu}$, we have the guarantee that any state run of $\Sigma \times \Sigma$ starting from $\mathcal{R}_0$ never reaches the unsafe region $\mathcal{R}_u$, i.e., $(\mathbf{x}_{x_0,\nu}(t), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(t)) \cap \mathcal{R}_u = \emptyset$, $\forall t \in \mathbb{N}$. This simply implies the satisfaction of $\|h(\mathbf{x}_{x_0,\nu}(t)) - h(\mathbf{x}_{\hat{x}_0,\hat{\nu}}(t))\| \leq \delta$, $\forall t \in \mathbb{N}$. Since $x_0 \in X_0 \cap X_S$ and $\mathbf{x}_{x_0,\nu}$ are arbitrarily chosen, we conclude that Σ is δ-approximate initial-state opaque. □

5.4.2 Verifying Lack of Approximate Initial-State Opacity

We presented in the previous subsection a sufficient condition for verifying approximate initial-state opacity based on a notion of barrier certificates. In particular, if one can find an ACBC satisfying conditions (5.3.1)-(5.3.3) with sets $\mathcal{R}_0$ and $\mathcal{R}_u$ defined as in (5.4.1)-(5.4.2), which ensures a safety property for the augmented system $\Sigma \times \Sigma$, then system Σ is shown to be approximate initial-state opaque. However, failing to find such an ACBC does not necessarily imply that the system is not opaque. Motivated by this, in this subsection, we aim at presenting a sufficient condition to verify the lack of

approximate initial-state opacity of a dt-CS Σ. Inspired by the duality between safety and reachability, our method is based on constructing another type of ACBC ensuring a reachability property for $\Sigma \times \Sigma$ as in Proposition 5.3.2.

The following theorem shows that the ACBC as in Proposition 5.3.2 can be used for verifying the lack of approximate initial-state opacity of dt-CSs.

Theorem 5.4.2. *Consider a dt-CS Σ as in Definition 2.3.1. Suppose there exists a continuous function $V : X \times X \to \mathbb{R}$ satisfying* (5.3.4)-(5.3.6) *in Proposition 5.3.2 with sets $\mathcal{R}_0, \mathcal{R}_u$ given in* (5.4.1)-(5.4.2)*. Then, system Σ is not δ-approximate initial-state opaque.*

Proof. First note that from Definition 3.3.1, system Σ is not δ-approximate initial-state opaque if there exists a state run $\mathbf{x}_{x_0,\nu}$ with $x_0 \in X_0 \cap X_S$, such that for any other state runs $\mathbf{x}_{\hat{x}_0,\hat{\nu}}$ starting from a non-secret initial condition $\hat{x}_0 \in X_0 \setminus X_S$, $\max_{i\in[0;n]} \|h(x_i) - h(\hat{x}_i)\| > \delta$ holds. Now consider a function $V : X \times X \to \mathbb{R}$ and an input run ν satisfying (5.3.6). Then, by Proposition 5.3.2 and from (5.4.1)-(5.4.2), it follows that there must exist a secret state $x_0 \in X_0 \cap X_S$ and a state run $\mathbf{x}_{x_0,\nu}$ under input run ν, such that for any trajectory $\mathbf{x}_{\hat{x}_0,\hat{\nu}}$ originated from any non-secret initial condition $\hat{x}_0 \in X_0 \setminus X_S$, the trajectories $(\mathbf{x}_{x_0,\nu}(t), \mathbf{x}_{\hat{x}_0,\hat{\nu}}(t))$ will eventually reach $\mathcal{R}_u$ in finite time, where $\|h(\mathbf{x}_{x_0,\nu}(t)) - h(\mathbf{x}_{\hat{x}_0,\hat{\nu}}(t))\| > \delta$. Therefore, for the state run $\mathbf{x}_{x_0,\nu}(t)$, there does not exist a state run starting from a non-secret initial state that generates similar output trajectories. Thus, δ-approximate initial-state opacity is violated. □

Remark 5.4.3. *We remark that the universal quantifier in* (5.3.4) *is not necessary to show the lack of approximate initial-state opacity. In fact, according to the duality of safety and reachability, the existence of one trajectory that starts from the initial region $\mathcal{R}_0$ and eventually enters into the unsafe region $\mathcal{R}_u$ is enough to show the lack of opacity. Therefore, one can relax the universal quantifier to an existential one by modifying the definition of barrier certificates, together with the corresponding initial and unsafe regions, at the cost of having a much more complex structure. However, it is difficult to formulate such a function and the corresponding set constraints to sum-of-squares programs (c.f. Section 5.5), and thus, is out of the scope of this thesis.*

5.4.3 Verifying Approximate Infinite-Step Opacity

In the last section, we showed how to use certain types of barrier certificates for the verification of (the lack of) initial-state opacity for discrete-time control systems. In order to verify infinite-step opacity using the above-defined barrier certificates as in Propositions 5.3.1 and 5.3.2, the sets of interest $\mathcal{R}_0, \mathcal{R}_u$ need to be redefined in a specific way to capture the initial and secret information of system Σ. In particular,

we define sets of initial states $\mathcal{R}_0$ and unsafe states $\mathcal{R}_u$ as:

$$\begin{aligned}\mathcal{R}_0 &= \{(x,\hat{x}) \in X_0 \times X_0 : x \notin X_S, ||h(x) - h(\hat{x})|| \leq \delta\} \cup \\ &\quad \{(x,\hat{x}) \in X_0 \times X_0 : x \in X_S, \hat{x} \notin X_S, ||h(x) - h(\hat{x})|| \leq \delta\}, \\ \mathcal{R}_u &= \{(x,\hat{x}) \in X \times X : x \in X_S, \hat{x} \in X_S\} \cup \\ &\quad \{(x,\hat{x}) \in X \times X : x \in X_S, \hat{x} \notin X_S, ||h(x) - h(\hat{x})|| > \delta\} \cup \\ &\quad \{(x,\hat{x}) \in X \times X : x \notin X_S, \hat{x} \in X, ||h(x) - h(\hat{x})|| > \delta\},\end{aligned} \tag{5.4.3}$$

where $\delta \in \mathbb{R}_{\geq 0}$ denotes the measurement precision of the outside intruder as introduced in Definition 3.3.1.

Remark 5.4.4. *The intuitions of the above definition for sets $\mathcal{R}_0$ and $\mathcal{R}_u$ are explained as follows. The unsafe set $\mathcal{R}_u$ as in* (5.4.3) *is defined as the union of three sets, where each set captures a certain scenario which violates approximate infinite-step opacity. The first case happens when both x and $\hat{x}$ belong to the set X_S. When the system's state x belongs to the secret set, opacity requires that $\hat{x}$ is not in this set, so that the desired alternative trajectory exists. Second case happens if x belongs to X_S, and $\hat{x}$ belongs to $X \setminus X_S$, but they are not δ close. This makes the two system's trajectories distinguishable from the intruder point of view. Third case happens if x belongs to $X \setminus X_S$, and $\hat{x}$ belongs to X, and they are not δ-close. In this case, since x does not belong to the secret set, we do not require $\hat{x}$ to belong to a certain subset of X. However, if the distance between the two trajectories exceeds δ, they would be distinguished by the intruder. Similarly, to define the initial set, we also consider possible initial conditions which the system can start from. First case is when x_0 belongs to $X_0 \setminus X_S$, $\hat{x}_0$ belongs to X_0, and they are δ close. This conveys if the system's initial condition is not secret, all we need for ensuring opacity of the system is to keep the trajectories δ-close. However, if the initial condition is secret, we require the alternative trajectory $\hat{x}$ to remain δ-close. This means x belongs to X_S, $\hat{x}$ belongs to $X_0 \setminus X_S$, and they are δ-close. Finally, we note that the sets defined to form $\mathcal{R}_0$ and $\mathcal{R}_u$ do not intersect.*

Now, we are ready to introduce the next theorem, which states the usefulness of the barrier certificates for verifying approximate infinite-step opacity of discrete-time control systems.

Theorem 5.4.5. *Consider a control system Σ as in Definition 2.3.1 and its associated augmented system $\Sigma \times \Sigma$. Suppose that there exists a function $B : X \times X \to \mathbb{R}_{\geq 0}$ satisfying conditions* (5.3.1)-(5.3.3) *in Proposition 5.3.1 with sets $\mathcal{R}_0$ and $\mathcal{R}_u$ defined as in* (5.4.3)*. Then, system Σ is δ-approximate infinite-step opaque.*

Proof. Let us first mention that, by applying the result from Proposition 5.3.1, the existence of a barrier certificate B ensures a safety property for the augmented system $\Sigma \times \Sigma$. That is, for any initial condition $(x_0, \hat{x}_0) \in \mathcal{R}_0$, and any input run ν, there exists an input run $\hat{\nu}$ such that $(\mathbf{x}_{x_0,\nu}, \mathbf{x}_{x_0,\nu}) \cap \mathcal{R}_u = \emptyset$.

Now, let the set of initial conditions $\mathcal{R}_0$ and unsafe states $\mathcal{R}_u$ be as defined in (5.4.3). Consider an arbitrary initial state x_0, an input sequence ν and the corresponding state run $\mathbf{x}_{x_0,\nu} = \{x_0, \ldots, x_n\}$ in Σ such that $x_k \in X_S$ for some $k \in \{0, \ldots, n\}$. We consider the following two cases:

If $k = 0$, then we have $x_0 \in X_S$. By the assumption that $\{x \in X_0 : \|h(x) - h(x_0)\| \leq \delta\} \not\subseteq X_S$ for any $x_0 \in X_0$, we know that there exists $\hat{x}_0 \in X \setminus X_S$ such that $\|h(x_0) - h(\hat{x}_0)\| \leq \delta$. Consider the augmented initial state $(x_0, \hat{x}_0)$, it can be readily verified that $(x_0, \hat{x}_0) \in \mathcal{R}_0$, where set $\mathcal{R}_0$ is as defined in (5.4.3). Then, as a consequence of the safety property of $\Sigma \times \Sigma$ (which is guaranteed from the existence of a barrier certificate B), we get that there exists an input run $\hat{\nu}$ such that the state run $(\mathbf{x}_{x_0,\nu}, \mathbf{x}_{\hat{x}_0,\hat{\nu}})$ of the augmented system $\Sigma \times \Sigma$ never reaches the unsafe set $\mathcal{R}_u$, i.e., $(\mathbf{x}_{x_0,\nu}, \mathbf{x}_{\hat{x}_0,\hat{\nu}}) \cap \mathcal{R}_u = \emptyset$. By the structure of $\mathcal{R}_u$, this implies the satisfaction of $\|h(\mathbf{x}_{x_0,\nu}(t)) - h(\mathbf{x}_{\hat{x}_0,\hat{\nu}}(t))\| \leq \delta$, for all $t \in \mathbb{N}$ (cf. Remark 5.4.4 for more intuitions on the structure of set $\mathcal{R}_u$).

If $k \geq 1$, then we have $x_0 \in X_0 \setminus X_S$. Again, we get by assumption that there exists $\hat{x}_0 \in X \setminus X_S$ such that $\|h(x_0) - h(\hat{x}_0)\| \leq \delta$. One can verify that the augmented initial state $(x_0, \hat{x}_0)$ also belongs to the set $\mathcal{R}_0$ as defined in (5.4.3). Again, by utilizing the safety property of $\Sigma \times \Sigma$, there exists an input run $\hat{\nu}$ such that the state run $(\mathbf{x}_{x_0,\nu}, \mathbf{x}_{\hat{x}_0,\hat{\nu}})$ of the augmented system $\Sigma \times \Sigma$ never reaches the unsafe set $\mathcal{R}_u$. Given that $x_k \in X_S$ and by further leveraging the structure of $\mathcal{R}_u$, it follows that $\mathbf{x}_{\hat{x}_0,\hat{\nu}}(t) \in X \setminus X_S$ and $\|h(\mathbf{x}_{x_0,\nu}(t)) - h(\mathbf{x}_{\hat{x}_0,\hat{\nu}}(t))\| \leq \delta$, for all $t \in \mathbb{N}$ (cf. Remark 5.4.4 for more intuitions on the structure of set $\mathcal{R}_u$).

Since the state run $\mathbf{x}_{x_0,\nu} = \{x_0, \ldots, x_n\}$ in Σ and index k are arbitrary, we can conclude that system Σ is δ-approximate infinite-step opaque. □

5.4.4 Verifying Lack of Approximate Infinite-Step Opacity

In the last subsection, we developed a sufficient condition for verifying approximate infinite-step opacity based on a notion of barrier certificates. Again, failure in finding such a barrier certificate does not imply the lack of opacity. Next, we introduce by the following proposition a sufficient condition for the the lack of approximate infinite-step opacity of Σ by searching for a barrier certificate which ensures a reachability property for the augmented system $\Sigma \times \Sigma$.

Proposition 5.4.6. *Consider a control system Σ as in Definition 2.3.1 and its associated augmented system $\Sigma \times \Sigma$. Suppose that there exists a function $V : X \times X \to \mathbb{R}_{\geq 0}$ satisfying conditions* (5.3.4)-(5.3.6) *in Proposition 5.3.2 with sets $\mathcal{R}_0$ and $\mathcal{R}_u$ defined as in* (5.4.3)*. Then, system Σ is not δ-approximate infinite-step opaque.*

Proof. The proof of this proposition follows by combining the result of Proposition 5.3.2 and Theorem 5.4.2. However, we should note that the definitions of the sets $\mathcal{R}_0$ and $\mathcal{R}_u$ are different in order to capture different notions of opacity. □

In the next section, we discuss how to leverage existing computational methods and software tools to compute $B(x, \hat{x})$ and $V(x, \hat{x})$ in Propositions 5.3.1 and 5.3.2, respectively.

5.5 Computation of Barrier Certificates using Sum-of-Squares Technique

In the previous sections, we presented sufficient conditions for verifying (resp. the lack of) approximate opacity of discrete-time control systems by searching for barrier certificates satisfying inequalities (resp. (5.3.4)-(5.3.6)) (5.3.1)-(5.3.3). For systems with polynomial transition functions and semi-algebraic sets (i.e., described by polynomial equalities and inequalities) X_0, X_S, and X, an efficient computational method based on sum-of-squares (SOS) programming can be utilized to search for polynomial barrier certificates.

Assumption 5.5.1. *A discrete-time control system Σ as in Definition 2.3.1 has continuous state set $X \subseteq R^n$ and continuous input set $U \subseteq \mathbb{R}^m$. Its transition function $f : X \times U \to X$ is polynomial in variables x and u, and output map h is polynomial in variable x.*

In the next lemma, we translate the conditions in Proposition 5.3.1 to SOS constraints.

Lemma 5.5.2. *Suppose Assumption 5.5.1 holds and sets $\mathcal{R}_0$, $\mathcal{R}_u$, $\mathcal{R}$, and U can be defined as $\mathcal{R}_0 = \{(x, \hat{x}) \in \mathbb{R}^n \times \mathbb{R}^n \mid g_0(x, \hat{x}) \geq 0\}$, $\mathcal{R}_u = \{(x, \hat{x}) \in \mathbb{R}^n \times \mathbb{R}^n \mid g_u(x, \hat{x}) \geq 0\}$, $\mathcal{R} = \{(x, \hat{x}) \in \mathbb{R}^n \times \mathbb{R}^n \mid g(x, \hat{x}) \geq 0\}$, $U = \{u \in \mathbb{R}^m \mid g_c(u) \geq 0\}$, where the inequalities are defined element-wise, and g_0, g_u, g, g_c are vectors of some polynomial functions. Suppose there exists a polynomial function $B(x, \hat{x})$, polynomials $p_{\hat{u}_i}(x, \hat{x}, u)$ corresponding to the i^{th} component of $\hat{u} = [\hat{u}_1; \hat{u}_2; \ldots; \hat{u}_m] \in U \subseteq \mathbb{R}^m$, and vectors of SOS polynomials $\lambda_0, \lambda_u, \lambda, \lambda_c$ of appropriate size such that the following expressions are SOS polynomials:*

$$-B(x, \hat{x}) - \lambda_0^\top(x, \hat{x}) g_0(x, \hat{x}) + \underline{\epsilon}, \tag{5.5.1}$$

$$B(x, \hat{x}) - \lambda_u^\top(x, \hat{x}) g_u(x, \hat{x}) - \overline{\epsilon}, \tag{5.5.2}$$

$$\begin{aligned} -B(f(x,u), f(\hat{x}, \hat{u})) + B(x, \hat{x}) - \lambda^\top(x, \hat{x}) g(x, \hat{x}) \\ - \sum_{i=1}^{m} (\hat{u}_i - p_{\hat{u}_i}(x, \hat{x}, u)) - \lambda_c^\top(u) g_c(u), \end{aligned} \tag{5.5.3}$$

where $\underline{\epsilon}, \overline{\epsilon} \in \mathbb{R}_{\geq 0}$ are some constants with $\overline{\epsilon} > \underline{\epsilon}$. Then, $B(x, \hat{x})$ satisfies conditions (5.3.1)-(5.3.3) *and $\hat{u} = [\hat{u}_1; \hat{u}_2; \ldots; \hat{u}_m]$, where $\hat{u}_i = p_{\hat{u}_i}(x, \hat{x}, u)$, $\forall i \in [1; m]$, is a control policy satisfying* (5.3.3).

We omit the proof of Lemma 5.5.2, since it follows the general methods for converting set constraints conditions to SOS programs with *Positivstellensatz* conditions, see [138] for details. Similarly, we convert the conditions of Proposition 5.3.2 to SOS constraints as well.

Lemma 5.5.3. *Suppose Assumption 5.5.1 holds and $X \subset \mathbb{R}^n$ is a bounded set. Suppose the regions of interest in Proposition 5.3.2 can be defined as $\mathcal{R}_0 = \{(x, \hat{x}) \in \mathbb{R}^n \times \mathbb{R}^n \mid$*

$g_0(x,\hat{x}) \geq 0\}$, $\partial\mathcal{R} \setminus \partial\mathcal{R}_u = \{(x,\hat{x}) \in \mathbb{R}^n \times \mathbb{R}^n \mid g_u(x,\hat{x}) \geq 0\}$, $\overline{(\mathcal{R} \setminus \mathcal{R}_u)} = \{(x,\hat{x}) \in \mathbb{R}^n \times \mathbb{R}^n \mid g(x,\hat{x}) \geq 0\}$, $U = \{\hat{u} \in \mathbb{R}^m \mid g_c(\hat{u}) \geq 0\}$, *where the inequalities are defined element-wise, and* g_0, g_u, g, g_c *are vectors of some polynomial functions. Suppose there exists a polynomial function* $V(x,\hat{x})$*, polynomials* $p_{u_i}(x,\hat{x},\hat{u})$ *corresponding to the* i^{th} *component of* $u = [u_1; u_2; \ldots; u_m] \in U \subseteq \mathbb{R}^m$*, and vectors of SOS polynomials* $\lambda_0, \lambda_u, \lambda, \lambda_c$ *of appropriate size such that the following expressions are SOS polynomials:*

$$-V(x,\hat{x}) - \lambda_0^\top(x,\hat{x})g_0(x,\hat{x}), \tag{5.5.4}$$

$$V(x,\hat{x}) - \lambda_u^\top(x,\hat{x})g_u(x,\hat{x}) - \varepsilon, \tag{5.5.5}$$

$$-V(f(x,u),f(\hat{x},\hat{u})) + V(x,\hat{x}) - \lambda^\top(x,\hat{x})g(x,\hat{x}) - \sum_{i=1}^{m}(u_i - p_{u_i}(x,\hat{x},\hat{u})) - \lambda_c^\top(\hat{u})g_c(\hat{u}) - \varepsilon, \tag{5.5.6}$$

where ε *is a small positive number. Then,* $V(x,\hat{x})$ *satisfies conditions* (5.3.4)-(5.3.6) *and* $u = [u_1; u_2; \ldots; u_m]$*, where* $u_i = p_{u_i}(x,\hat{x},\hat{u})$*,* $\forall i \in [1; m]$*, is a control policy satisfying* (5.3.6)*.*

Note that a small tolerance ε in (5.5.5) and (5.5.6) is needed to ensure positivity of polynomials as required in (5.3.5) and (5.3.6).

Remark 5.5.4. *As seen in Lemmas 5.5.2 and 5.5.3, in order to search for polynomial barrier certificates by means of SOS programming, it is required that regions* $\mathcal{R}_0$*,* $\mathcal{R}_u$*,* $\partial\mathcal{R} \setminus \partial\mathcal{R}_u$*,* $\overline{(\mathcal{R} \setminus \mathcal{R}_u)}$ *are semi-algebraic sets. We highlight that having a system* Σ *with semi-algebraic sets* X_0*,* X_S*, and* X *is enough to ensure that all these regions are semi-algebraic. In particular, as a consequence of Tarski-Seidenberg principle [187], the class of all semi-algebraic sets is closed under finite unions, intersections, taking complement, and Cartesian product. The boundary, the interior, and the closure of a semi-algebraic set are also semi-algebraic. Additionally, given the polynomial output map* h*, the set of states satisfying* $\|h(x) - h(\hat{x})\| \leq \delta$ *is equivalent to the one satisfying* $(h(x) - h(\hat{x}))^\top(h(x) - h(\hat{x})) \leq \delta^2$*, which is again a semi-algebraic set. See [37] for details.*

One can leverage existing computational toolboxes such as `SOSTOOLS` [138] together with semidefinite programming solvers such as `SeDuMi` [179] to compute polynomial barrier certificates satisfying (5.5.1)-(5.5.3) or (5.5.4)-(5.5.6).

Remark 5.5.5. *By formulating conditions* (5.3.1)-(5.3.3) *(resp.* (5.3.4)-(5.3.6)*) as a satisfiability problem, one can alternatively search for parametric control barrier certificates using an iterative program synthesis framework, called Counter-Example-Guided Inductive Synthesis (CEGIS), with the help of Satisfiability Modulo Theories (SMT) solvers such as Z3 [41] and dReal [48]; see, e.g., [76] for more details. We also refer interested readers to the recent work [140], where machine learning techniques were exploited for the construction of barrier certificates.*

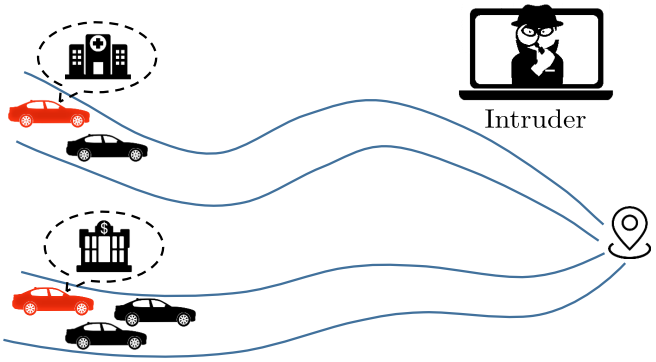


Figure 5.3: Plausible deniability of a vehicle in terms of its initial conditions. The blue lines roughly indicate the intruder's insufficient observation precision.

5.6 Case Studies

Here, we provide two examples to illustrate how one can utilize the theoretical results obtained in the previous sections for the verification of (the lack of) approximate initial-state opacity.

5.6.1 Verifying Approximate Initial-State Opacity on a Vehicle Model

In this example, we consider an autonomous vehicle moving on a single lane road, whose state variable is defined as $x = [x_1; x_2]$, with x_1 being its absolute position (in the road frame) and x_2 being its absolute velocity. The discrete-time dynamics of the vehicle is modeled as:

$$\begin{aligned} \begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} &= \begin{bmatrix} 1 & \Delta\tau \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} \Delta\tau^2/2 \\ \Delta\tau \end{bmatrix} u(t), \\ y(t) &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \end{aligned} \tag{5.6.1}$$

where u is the control input (acceleration) and $\Delta\tau$ is the sampling time. The output is assumed to be the position of the vehicle on the road. Let us first briefly explain the motivation behind this example; see Figure 5.3. Suppose the initial locations of the vehicle contain critical information which is needed to be kept secret, e.g., the vehicle might be a cash transit van that aims at transferring money initially from a bank to an ATM machine, or a patient who initially visited a hospital but unwilling to reveal personal information to others. It is implicitly assumed that there is a malicious intruder who is observing the behavior of the vehicle remotely intending to carry out an attack. Therefore, it is in the interest of the system to verify whether it maintains plausible deniability for secret initial conditions where some confidential assignment is executed. This problem can be formulated as a δ-approximate initial-state opacity problem, where $\delta \geq 0$ captures the security-guarantee level in terms of the measurement precision of the intruder. Now consider system (5.6.1) with state space $X = [0, 10] \times$

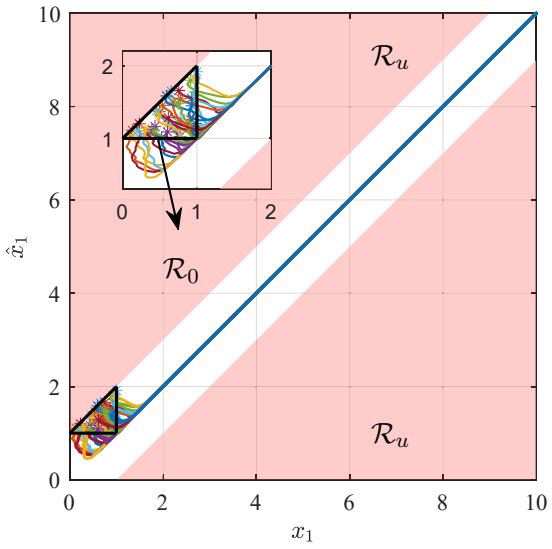


Figure 5.4: Trajectories of $\Sigma \times \Sigma$ projected on the position plane starting from initial region $\mathcal{R}_0$ (represented by the black triangle). The regions in red are the unsafe set $\mathcal{R}_u$.

$[0, 0.1]$, initial set $X_0 = [0, 10] \times \{0\}$, secret set $X_S = [0, 1] \times [0, 0.1]$, input set $U = [-0.05, 0.05]$ and sampling time $\Delta\tau = 1$. Consider the augmented system $\Sigma \times \Sigma$. Accordingly, the regions of interest in (5.4.1) and (5.4.2) are $\mathcal{R}_0 = \{[x_1; x_2] \in [0, 1] \times \{0\}, [\hat{x}_1; \hat{x}_2] \in [1, 10] \times \{0\} \mid (x_1 - \hat{x}_1)^2 \leq \delta^2\}$, and $\mathcal{R}_u = \{(x, \hat{x}) \in X \times X \mid (x_1 - \hat{x}_1)^2 \geq \delta^2 + \epsilon\}$. Note that a small positive number ϵ is needed to certify positivity of the obtained polynomials using SOS programming. Now, we set the threshold parameter to be $\delta = 1$ and search for barrier certificates by solving sum-of-squares programs with the help of `SOSTOOLS` and `SeDuMi` tools as described in Section 5.5. Using Lemma 5.5.2, we obtained a polynomial ACBC of degree 2 satisfying (5.5.1)-(5.5.3) with $\underline{\epsilon} = 1$, $\overline{\epsilon} = 1.001$ and a tolerance $\epsilon = 0.01$ as follows

$$\begin{aligned} B(x, \hat{x}) &= 0.9227x_1^2 + 0.2348x_2^2 + 0.9227\hat{x}_1^2 + 0.2348\hat{x}_2^2 \\ &+ 0.006x_1x_2 - 0.006\hat{x}_1x_2 - 0.006x_1\hat{x}_2 - 0.006\hat{x}_1\hat{x}_2 \\ &- 0.4696x_2\hat{x}_2 - 1.845x_1\hat{x}_1 - 0.0002\hat{x}_1 + 0.0728, \end{aligned}$$

and the corresponding control policy is $\hat{u}(x, \hat{x}, u) = 0.8x_1 - 0.8x_2 + 1.5\hat{x}_1 - 1.5\hat{x}_2 + u$. Therefore, we conclude that Σ is 1-approximate initial-state opaque. Particularly, for every trajectory starting from a secret state, there always exists at least one alternative trajectory originated from a non-secret state which are indistinguishable for an intruder with measurement precision δ. Figure 5.4 shows the projection of a few state trajectories on the position plane of the augmented system $\Sigma \times \Sigma$, starting from randomly generated initial conditions in $\mathcal{R}_0$ under control policy $\hat{u}$ with u taking values in U. It is seen that any trajectory starting from $\mathcal{R}_0$ does not reach the unsafe region $\mathcal{R}_u$ as time increases. We further notice that $\delta = 1$ is the smallest threshold for which we are able to find a

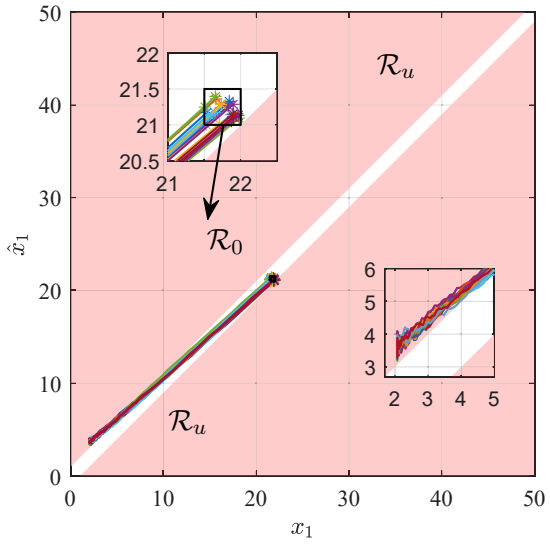


Figure 5.5: Trajectories of $\Sigma \times \Sigma$ projected on the first-room plane starting from initial region $\mathcal{R}_0$ (represented by the black rectangle). The regions in red are the unsafe set $\mathcal{R}_u$.

barrier certificate ensuring approximate initial-state opacity. For a smaller value of δ, approximate initial-state opacity is immediately violated at the initial condition.

5.6.2 Verifying Lack of Approximate Initial-State Opacity on a Room Temperature Model

In this example, we showcase the use of an ACBC in verifying the lack of opacity in a two-room temperature model by Proposition 5.3.2. The model is borrowed from [127]. The evolution of the temperature $\mathbf{T}(\cdot)$ of 2 rooms is described by the discrete-time model:

$$\Sigma : \begin{cases} \mathbf{T}(k+1) = & A\mathbf{T}(k) + \alpha_h T_h \nu(k) + \alpha_e T_e, \\ \mathbf{y}(k) = & h(\mathbf{T}(k)), \end{cases} \tag{5.6.2}$$

where $A \in \mathbb{R}^{2\times 2}$ is a matrix with elements $\{A\}_{ii} = (1-2\alpha-\alpha_e-\alpha_h\nu_i)$, $\{A\}_{12} = \{A\}_{21} = \alpha$, $\mathbf{T}(k) = [\mathbf{T}_1(k);\ \mathbf{T}_2(k)]$, $T_e = [T_{e1}; T_{e2}]$, $\nu(k) = [\nu_1(k); \nu_2(k)]$, where $\nu_i(k) \in [0,1]$, $\forall i \in [1;2]$, represents the ratio of the heater valve being open in room i. The output of the network is assumed to be the temperature of the second room: $h(\mathbf{T}(k)) = \mathbf{T}_2(k)$. Parameters $\alpha = 0.05$, $\alpha_e = 0.008$, and $\alpha_h = 0.0036$ are heat exchange coefficients, $T_{e1} = T_{e2} = -1\,^\circ C$ is the external temperature, and $T_h = 50\,^\circ C$ is the heater temperature. The regions of interest in this example are $X = [0,50]^2$, $X_0 = [21,22]^2$, and $X_S = [21.5,50] \times [0,50]$. Specifically, the secret of the network is whether the first room has a temperature initially higher than $21.5\,^\circ C$ (which may indicate activities with people gathering in that room). The intruder wants to infer the initial temperature of the first room by monitoring the temperature variation of the last room and using the knowledge of the system model. Now the objective is to verify if the system is able

to keep this secret in the presence of a malicious intruder with measurement precision $\delta = 1$. In this example, a degree bound of 8 is imposed on B and V. First, by means of SOSTOOLS, we failed to find a function $B(x, \hat{x})$ satisfying (5.5.1)-(5.5.3) in Lemma 5.5.2. Then, we compute a function $V(x, \hat{x})$ as in Lemma 5.5.3 to see if the system is lacking the approximate initial-state opacity. In this case, the regions considered in Lemma 5.5.3 are

$$\begin{aligned}
&\mathcal{R}_0 = \{\mathbf{T} \in [21.5, 22] \times [21, 22], \hat{\mathbf{T}} \in [21, 21.5] \times [21, 22]\},\\
&\partial\mathcal{R} \setminus \partial\mathcal{R}_u = \{(\mathbf{T}, \hat{\mathbf{T}}) \in \mathcal{R} \mid (\mathbf{T}_1, \hat{\mathbf{T}}_1) \in \mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_3 \cup \mathcal{R}_4\},\\
&\overline{(\mathcal{R} \setminus \mathcal{R}_u)} = \{(\mathbf{T}, \hat{\mathbf{T}}) \in \mathcal{R} \mid (\mathbf{T}_1 - \hat{\mathbf{T}}_1)^2 \leq \delta^2\},
\end{aligned}$$

where $\mathcal{R} = X \times X$, $\mathcal{R}_1 = [0, \delta] \times \{0\}$, $\mathcal{R}_2 = \{0\} \times [0, \delta]$, $\mathcal{R}_3 = \{50\} \times [50 - \delta, 50]$, $\mathcal{R}_4 = [50 - \delta, 50] \times \{50\}$. With the aid of SOSTOOLS and SeDuMi, we obtained a polynomial barrier certificate of degree 6 satisfying (5.5.4)-(5.5.6) with a tolerance $\varepsilon = 0.01$ and control policy $\nu(k) = [0; 0], \forall k \in \mathbb{N}_{\geq 0}$. The system is thus lacking 1-approximate initial-state opacity. This means that for each state run starting from a secret initial state in Σ under ν, all trajectories from non-secret states will eventually deviate from the former ones in the sense of generating different outputs (captured by δ). Once the intruder sees these trajectories, it is certain that the system was initiated from a secret state. Figure 5.5 shows trajectories of $\Sigma \times \Sigma$ from $\mathcal{R}_0$ under control sequence ν with $\hat{\nu}$ taking values in U. The trajectories eventually reach $\mathcal{R}_u$ in finite time.

5.7 Discussion and Future Work

In this chapter, we proposed a discretization-free framework for opacity verification of discrete-time control systems. A pair of augmented control barrier certificates were defined for the analysis of different types of approximate opacity, which were constructed over an augmented system that is the product of a control system and itself. While both barrier certificates only serve as sufficient conditions, they can be utilized in reverse directions in the sense that one ensures approximate opacity, and the other one shows the lack of approximate opacity of the control system. We showed that the computation of the barrier certificates can be carried out by some SOS programming. Numerical case studies were conducted to illustrate the effectiveness of the proposed results.

Future Work As we have already mentioned, there are very few results for abstraction-free opacity synthesis. One important direction is to extend the barrier-certificates techniques from opacity verification to opacity synthesis. To this end, one may borrow the idea of control barrier functions [7, 168] that generalizes the idea of barrier certificates to control systems by explicitly taking the effect of control choices into account. Moreover, one can further develop a secure-by-construction scheme for synthesizing controllers to enforce safety (or more general mission requirements) and security properties simultaneously over control systems. This can be achieved by establishing a

bridge between the desired safety property and security property by leveraging notions of (augmented) control barrier functions.

6 Modular Verification of Opacity for Large-scale Interconnected Systems

6.1 Introduction

In the previous sections, we presented various abstraction-based and discretization-free approaches in verifying opacity for CPS. Though promising, when confronted with large-scale interconnected systems, the aforementioned results will become computationally intractable. This prevents current techniques from providing automated verification or synthesis for large-scale interconnected CPS. This is not just a theoretical concern, many safety-critical applications, such as traffic network, automated highway driving, building management systems, power networks, air traffic management, uninhabited aerial vehicles, and so on, consist of many subsystems interacting with each other. One way to address the inherent difficulty in analyzing or controlling complex, large-scale, interconnected systems, is to apply a "divide and conquer" strategy, namely, compositional (modular) approaches.

6.1.1 Related Literature

As we have discussed in Chapter 4, opacity-preserving finite abstractions and simulation relations serve as a bridge between continuous-space CPS and existing verification or synthesis algorithms for opacity developed in DES community. Although they are shown to be a useful tool, a challenge lies in scaling the approach for large-scale systems. Typically, the proposed techniques reported in Chapter 4 take a monolithic view of systems where abstraction, verification, and synthesis are performed for the entire system. This monolithic view interacts poorly with the construction of finite abstractions and in general suffers from the so-called the *curse of dimensionality*: the complexity of constructing the abstraction grows exponentially with the state dimension of the model. Different compositional approaches have been proposed in the literature to overcome this challenge in dealing with large-scale CPS. The two most commonly used schemes are based on: 1) assume-guarantee contracts [88, 169, 176, 111] which are originally introduced in the computer science literature and 2) the input-output properties of the system, including those expressed as small-gain [159, 88, 146] or dissipativity properties [215, 182] which are originally introduced in the control theory literature. Here, the overall large-scale systems are usually seen as interconnections of smaller (reasonably sized) components, i.e., subsystems. Subsequently, the analysis and the design of the overall system is reduced to those of the subsystems.

In the past decades, many potential compositionality results have been proposed to tackle the acute computational bottlenecks in the analysis of safety properties for large-scale continuous-space systems [188, 146, 87, 89, 22, 159, 182, 184, 110, 98, 116]. Among the many existing compositionality results, recent works [146, 127, 89, 184, 185, 183, 110] proposed compositional techniques for constructing finite abstractions for networks of systems. The results in [146] first explored small-gain conditions for the compositional construction of *complete* finite abstractions for a network of discrete-time control systems. The results in [184] proposed a max-type small-gain type compositional condition which results in a finite abstraction with smaller approximation error. There are also other results in the literature [127, 89] which provide compositional construction of *sound* abstractions without imposing strong compositionality conditions. However, the aforementioned compositional schemes above are proposed for the sake of controller synthesis for temporal logic properties, and none of them are applicable to deal with security properties including opacity. Recently, a compositional framework is presented in [143] motivated by the computational complexity encountered in the analysis of a related property, called *critical observability*, for large-scale networks of finite state machines. A bisimulation equivalence is defined taking into account criticalities. In the context of analyzing security properties, compositional approaches have been explored only recently for modular verification and synthesis of DES in [164, 135, 132, 190, 204, 224].

6.1.2 Contributions

Motivated by the computational complexity issues appeared in the verification techniques in the previous chapters, here, we aim at providing a compositional framework to conquer this complexity challenge using a "divide and conquer" strategy.

In the first part of this chapter, we propose a modular opacity verification approach for interconnected control system using abstraction techniques. In particular, the result here is based on the compositional construction of opacity-preserving finite abstractions for networks of discrete-time nonlinear control systems. To this purpose, we first introduce new notions of opacity-preserving simulation functions for both subsystems and the entire networks tailored to the three basic notions of approximate opacity. Based on these notions, we propose a compositional scheme on the construction of abstractions for concrete networks. Rather than dealing with the original large-scale system, our compositional framework allows one to construct opacity-preserving abstractions locally using local opacity-preserving simulation functions, while providing the guarantee that the interconnection of local abstractions simulates the concrete network while preserving opacity across them. By exploiting the interconnection topology of the network, an algorithm is presented to orderly design local quantization parameters with the guarantee of obtaining an overall finite abstraction with any given desired precision.

In the second part of this chapter, we enlarge the class of systems to hybrid ones with switching signals. That is, we provide a compositional approach to verify approximate opacity of a network of switched systems using their finite abstractions. Note that if switched subsystems accept common incremental Lyapunov functions, our pro-

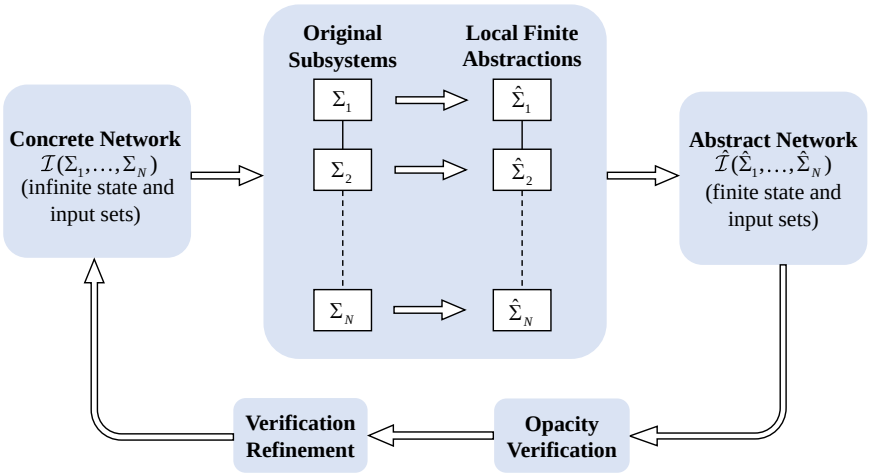


Figure 6.1: Compositional framework for opacity verification of networks of systems.

posed results here recover the ones presented in the first part of this chapter. Here, we consider two types of approximate opacity, i.e., approximate initial-state opacity and approximate current-state opacity. A new notion of approximate initial-state (resp. current-state) opacity-preserving simulation function (InitSOPSF, resp. CurSOPSF) is introduced as a system relation to characterize the closeness between two networks in terms of preservation of approximate initial-state (resp. current-state) opacity. We show that such an InitSOPSF (resp. CurSOPSF) can be established by composing certain local InitSOPSFs (resp. CurSOPSFs) which relates each switched subsystem to its local finite abstraction. Moreover, under some assumptions ensuring incremental input-to-state stability of discrete-time switched systems, an approach is provided to construct local finite abstractions along with the corresponding local InitSOPSFs (resp. CurSOPSFs) for subsystems. Then, we derive some small-gain type conditions, under which one can construct a finite abstraction of the concrete network of switched systems by interconnecting local finite abstractions of subsystems. Finally, one can verify opacity based on the constructed finite abstraction, and then refine the results back to the concrete network based on their opacity-preserving system relation. The proposed compositional abstraction-based opacity verification pipeline is depicted in Figure 6.1.

In the third part of this chapter, we develop a modular opacity verification approach by compositional construction of barrier certificates for large-scale interconnected systems. Notice that as presented in Section 5.4, barrier certificates can be leveraged as a useful alternative approach for the verification of opacity for CPS. Though promising, the computation of such types of barrier certificates is still an expensive problem, which may become intractable while dealing with large-scale interconnected systems. To this end, we define an *augmented system* by taking the product of an interconnected system with itself. Then, we construct barrier certificates for this augmented system *compositionally* by leveraging so-called local barrier certificates of augmented versions

of subsystems. The barrier certificate for the interconnected system is then constructed by composing those easier-to-compute local barrier certificates under some small-gain type conditions. We show that the existence of such barrier certificates is sufficient to ensure approximate opacity of the interconnected system.

6.2 An Abstraction-based Approach for Interconnected Control Systems

As mentioned earlier, while leveraging compositional approaches, the overall large-scale systems are usually seen as interconnections of smaller (reasonably sized) components, i.e., subsystems. Subsequently, the analysis and the design of the overall system is reduced to those of the subsystems. In this section, we first introduce the formal definitions of interconnected control systems, and then present our new notions of opacity-preserving simulation functions. The main compositionality result will be then provided based on a small-gain type condition.

6.2.1 Interconnected Control Systems

In the following, we denote a discrete-time control subsystem by a tuple $\Sigma_i = (X_i, X_{0_i}, X_{S_i}, U_i, W_i, f_i, Y_i, h_i)$. The formal definition of a control subsystem is similar to the one in Definition 2.3.1 but with two sets of inputs. In particular, $w_i \in W_i$ are termed as "internal" inputs which are used to describe the interaction between subsystems, and $u_i \in U_i$ are called "external" inputs served as interfaces for controllers. An interconnected control system composed of $N \in \mathbb{N}_{\geq 1}$ subsystems is iteself a discrete-time control system as in Definition 2.3.1, denoted by $\mathcal{I}(\Sigma_1, \ldots, \Sigma_N)$, subject to certain interconnection constraints.

6.2.1.1 Discrete-time Control Subsystems

In this section we study the class of discrete-time control subsystems of the following form.

Definition 6.2.1. *A discrete-time control subsystem Σ is defined by the tuple $\Sigma_i = (X_i, X_{0_i}, X_{S_i}, U_i, W_i, f_i, Y_i, h_i)$ where X_i, U_i, W_i and Y_i are the state, external input, internal input, and output set, respectively. We denote by $X_{0_i}, X_{S_i} \subseteq X_i$ the set of initial states and secret states, respectively. The set-valued map $f_i : X_i \times U_i \times W_i \rightrightarrows X_i$ is the state transition function, and $h_i : X_i \to Y_i$ is the output function. The discrete-time control subsystem Σ_i is described by difference inclusions of the form*

$$\Sigma_i : \begin{cases} \mathbf{x}_i(t+1) \in & f_i(\mathbf{x}_i(t), \nu_i(t), w_i(t)), \\ \mathbf{y}_i(t) = & h(\mathbf{x}_i(t)), \end{cases} \tag{6.2.1}$$

where $\mathbf{x}_i : \mathbb{N} \to X_i$, $\mathbf{y}_i : \mathbb{N} \to Y_i$, $\nu_i : \mathbb{N} \to U$, and $w_i : \mathbb{N} \to W_i$ are the state, output, external input, and internal input signals, respectively. System $\Sigma_i = (X_i, X_{0_i}, X_{S_i}, U_i, W_i, f_i, Y_i, h_i)$ is called deterministic if $\mathrm{card}(f_i(x_i, u_i, w_i)) \leq 1$ $\forall x_i \in X_i, \forall u_i \in U_i, \forall w_i \in W_i$,

and non-deterministic otherwise. System Σ_i is called finite if X_i, U_i, W_i are finite sets and infinite otherwise.

6.2.1.2 Discrete-time Interconnected Control Systems

Consider $N \in \mathbb{N}_{\geq 1}$ systems Σ_i as in Definition 6.2.1, $i \in [1; N]$. Assume internal inputs and output maps are partitioned as

$$w_i = [w_{i1}; \ldots; w_{i(i-1)}; w_{i(i+1)}; \ldots; w_{iN}], \tag{6.2.2}$$

$$h_i(x_i) = [h_{i1}(x_i); \ldots; h_{iN}(x_i)], \tag{6.2.3}$$

with $W_i = \prod_{j=1, j\neq i}^{N} W_{ij}$ and $Y_i = \prod_{j=1}^{N} Y_{ij}$, $w_{ij} \in W_{ij}$, $y_{ij} = h_{ij}(x_i) \in Y_{ij}$. The outputs y_{ii} are considered as external ones, whereas y_{ij} with $i \neq j$ are interpreted as internal ones to construct interconnections between subsystems. In the case that no connection exists between subsystems Σ_i and Σ_j, we simply have $h_{ij} \equiv 0$. Now, we are ready to provide a formal definition of interconnected dt-CS as follows. Note that an interconnected control system without internal inputs and outputs reduces to an discrete-time control system as in Definition 2.3.1.

Definition 6.2.2. *Consider $N \in \mathbb{N}_{\geq 1}$ discrete-time control subsystems $\Sigma_i = (X_i, X_{0_i}, X_{S_i}, U_i, W_i, f_i, Y_i, h_i)$, $i \in [1; N]$, with the input-output structure given in* (6.2.2)-(6.2.3)*. The* concrete *interconnected control system denoted by $\mathcal{I}(\Sigma_1, \ldots, \Sigma_N)$ is a tuple $\Sigma = (X, X_0, X_S, U, f, Y, h)$, where $X = \prod_{i=1}^{N} X_i, X_0 = \prod_{i=1}^{N} X_{0_i}, X_S = \prod_{i=1}^{N} X_{S_i}, U = \prod_{i=1}^{N} U_i$, $Y = \prod_{i=1}^{N} Y_{ii}$, $f(x, u) = \{[x'_1; \ldots; x'_N] |\ x'_i \in f_i(x_i, u_i, w_i), \forall i \in [1; N]\}$, $h(x) = [h_{11}(x_1); \ldots; h_{NN}(x_N)]$, subject to:*

$$y_{ji} = w_{ij}, Y_{ji} \subseteq W_{ij}, \forall i \in [1; N], j \neq i. \tag{6.2.4}$$

A finite *interconnected control system $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h})$, denoted by $\hat{\Sigma} = \hat{\mathcal{I}}(\hat{\Sigma}_1, \ldots, \hat{\Sigma}_N)$, is composed of $N \in \mathbb{N}_{\geq 1}$ finite discrete-time control subsystems $\hat{\Sigma}_i$, subject to:*

$$\|\hat{y}_{ji} - \hat{w}_{ij}\| \leq \phi_{ij}, [\hat{Y}_{ji}]_{\phi_{ij}} \subseteq \hat{W}_{ij}, \forall i \in [1; N], j \neq i, \tag{6.2.5}$$

where ϕ_{ij} is an internal input quantization parameter designed for constructing local finite abstractions (cf. Subsection 6.2.3.1).

An example of an interconnected system composed of two subsystems is depicted in Figure 6.2.

Remark 6.2.3. *Note that in the above definition, the interconnection constraint in* (6.2.4) *for the concrete network is different from that for the abstract network in* (6.2.5)*. For networks of finite abstractions, due to possibly different granularities of finite internal input sets $\hat{W}_{ij}$ and output sets $\hat{Y}_{ij}$, we introduce parameters ϕ_{ij} in* (6.2.5) *for having a well-posed interconnection.*

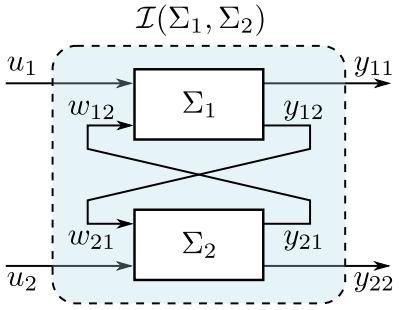


Figure 6.2: Feedback composition of two subsystems.

6.2.2 Opacity-Preserving Simulation Functions

In this section, we introduce new notions of approximate opacity-preserving simulation functions for both subsystems and interconnected systems, which will provide us the basis for using abstraction-based technique in verifying approximate opacity for large-scale interconnected systems.

6.2.2.1 Opacity-Preserving Simulation Functions

First, we introduce a new notion of initial-state opacity-preserving simulation functions.

Definition 6.2.4. *Consider interconnected dt-CS* $\Sigma = (X, X_0, X_S, U, f, Y, h)$ *and* $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h})$ *where* $\hat{Y} \subseteq Y$. *For* $\varpi \in \mathbb{R}_{\geq 0}$, *function* $\tilde{V} : X \times \hat{X} \to \mathbb{R}_{\geq 0}$ *is an* ϖ*-approximate initial-state opacity-preserving simulation function (*ϖ*-InitSOPSF) from* Σ *to* $\hat{\Sigma}$, *if there exists a function* $\alpha \in \mathcal{K}_\infty$ *s.t.*

1. *a)* $\forall x_0 \in X_0 \cap X_S$, $\exists \hat{x}_0 \in \hat{X}_0 \cap \hat{X}_S$, *s.t.* $\tilde{V}(x_0, \hat{x}_0) \leq \varpi$;
   b) $\forall \hat{x}_0 \in \hat{X}_0 \setminus \hat{X}_S$, $\exists x_0 \in X_0 \setminus X_S$, *s.t.* $\tilde{V}(x_0, \hat{x}_0) \leq \varpi$;
2. $\forall x \in X, \forall \hat{x} \in \hat{X}$, $\alpha(\|h(x) - \hat{h}(\hat{x})\|) \leq \tilde{V}(x, \hat{x})$;
3. $\forall x \in X, \forall \hat{x} \in \hat{X}$ *s.t.* $\tilde{V}(x, \hat{x}) \leq \varpi$, *the following hold:*
   a) $\forall u \in U$, $\forall x_d \in f(x, u)$, $\exists \hat{u} \in \hat{U}$, $\exists \hat{x}_d \in \hat{f}(\hat{x}, \hat{u})$, *s.t.* $\tilde{V}(x_d, \hat{x}_d) \leq \varpi$;
   b) $\forall \hat{u} \in \hat{U}$, $\forall \hat{x}_d \in \hat{f}(\hat{x}, \hat{u})$, $\exists u \in U$, $\exists x_d \in f(x, u)$, *s.t.* $\tilde{V}(x_d, \hat{x}_d) \leq \varpi$.

It is worth noting that the ϖ-InitSOPSF characterizes the distance between two systems in terms of the satisfaction of approximate opacity. This relation considers not only the dynamic, but also the secrets of the system. The usefulness of Definition 6.2.4 in terms of preservation of approximate opacity across related systems will be shown later in Proposition 6.2.8.

Next, we introduce a new notion of current-state opacity-preserving simulation functions.

Definition 6.2.5. *Consider interconnected dt-CS* $\Sigma = (X, X_0, X_S, U, f, Y, h)$ *and* $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h})$ *where* $\hat{Y} \subseteq Y$*. For* $\varpi \in \mathbb{R}_{\geq 0}$*, function* $\tilde{V} : X \times \hat{X} \to \mathbb{R}_{\geq 0}$ *is an* ϖ*-approximate current-state opacity-preserving simulation function (*ϖ*-CurSOPSF) from* Σ *to* $\hat{\Sigma}$*, if there exists a function* $\alpha \in \mathcal{K}_\infty$ *such that*

1. $\forall x_0 \in X_0$, $\exists \hat{x}_0 \in \hat{X}_0$, *s.t.* $\tilde{V}(x_0, \hat{x}_0) \leq \varpi$;
2. $\forall x \in X, \forall \hat{x} \in \hat{X}$, $\alpha(\|h(x) - \hat{h}(\hat{x})\|) \leq \tilde{V}(x, \hat{x})$;
3. $\forall x \in X$, $\forall \hat{x} \in \hat{X}$ *s.t.* $\tilde{V}(x, \hat{x}) \leq \varpi$, *the following hold:*
   a) $\forall u \in U$, $\forall x_d \in f(x, u)$, $\exists \hat{u} \in \hat{U}$, $\exists \hat{x}_d \in \hat{f}(\hat{x}, \hat{u})$, *s.t.* $\tilde{V}(x_d, \hat{x}_d) \leq \varpi$;
   b) $\forall u \in U$, $\forall x_d \in f(x, u)$ *s.t.* $x_d \in X_S$, $\exists \hat{u} \in \hat{U}$, $\exists \hat{x}_d \in \hat{f}(\hat{x}, \hat{u})$ *with* $\hat{x}_d \in \hat{X}_S$, *s.t.* $\tilde{V}(x_d, \hat{x}_d) \leq \varpi$;
   c) $\forall \hat{u} \in \hat{U}$, $\forall \hat{x}_d \in \hat{f}(\hat{x}, \hat{u})$, $\exists u \in U$, $\exists x_d \in f(x, u)$, *s.t.* $\tilde{V}(x_d, \hat{x}_d) \leq \varpi$;
   d) $\forall \hat{u} \in \hat{U}$, $\forall \hat{x}_d \in \hat{f}(\hat{x}, \hat{u})$ *s.t.* $\hat{x}_d \in \hat{X} \setminus \hat{X}_S$, $\exists u \in U$, $\exists x_d \in f(x, u)$ *with* $x_d \in X \setminus X_S$, *s.t.* $\tilde{V}(x_d, \hat{x}_d) \leq \varpi$.

Similarly, we introduce a new notion of infinite-step opacity-preserving simulation functions by combining the conditions of ϖ-InitSOPSF and ϖ-CurSOPSF.

Definition 6.2.6. *Consider interconnected dt-CS* $\Sigma = (X, X_0, X_S, U, f, Y, h)$ *and* $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h})$ *where* $\hat{Y} \subseteq Y$*. For* $\varpi \in \mathbb{R}_{\geq 0}$*, function* $\tilde{V} : X \times \hat{X} \to \mathbb{R}_{\geq 0}$ *is an* ϖ*-approximate infinite-step opacity-preserving simulation function (*ϖ*-InfSOPSF) from* Σ *to* $\hat{\Sigma}$*, if it is both an* ϖ*-InitSOPSF and an* ϖ*-CurSOPSF from* Σ *to* $\hat{\Sigma}$*.*

Note that if there exists an opacity-preserving simulation function from Σ to $\hat{\Sigma}$, and $\hat{\Sigma}$ is finite, $\hat{\Sigma}$ is called a finite abstraction of the concrete network Σ.

The next result shows that the existence of an ϖ-InitSOPSF (resp. CurSOPSF, InfSOPSF) for interconnected systems implies the existence of an ε-InitSOP (resp. CurSOP, InfSOP) simulation relation between them.

Proposition 6.2.7. *Consider interconnected dt-CS* $\Sigma = (X, X_0, X_S, U, f, Y, h)$ *and* $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h})$ *where* $\hat{Y} \subseteq Y$*. Assume* $\tilde{V}$ *is a* ϖ*-InitSOPSF (resp. CurSOPSF, InfSOPSF) from* Σ *to* $\hat{\Sigma}$ *with the corresponding function* $\alpha \in \mathcal{K}_\infty$ *as in Definitions 6.2.4, 6.2.5, and 6.2.6. Then, relation* $R \subseteq X \times \hat{X}$ *defined by*

$$R = \left\{(x, \hat{x}) \in X \times \hat{X} | \tilde{V}(x, \hat{x}) \leq \varpi\right\},$$

is an ε*-InitSOP (resp. CurSOP, InfSOP) simulation relation, defined as in Definition 4.3.1 (resp. Definition 4.3.6, Definition 4.3.8), from* Σ *to* $\hat{\Sigma}$ *with*

$$\varepsilon = \alpha^{-1}(\varpi). \tag{6.2.6}$$

Proof. Here, we prove the case for initial-state opacity-preserving simulation relation. The first condition in Definition 4.3.1 follows immediately from condition 1 in Definition 6.2.4, i.e. $\tilde{V}(x_0, \hat{x}_0) \leq \varpi$. Now, we show that $\forall (x, \hat{x}) \in R$: $\|h(x) - \hat{h}(\hat{x})\| \leq \varepsilon$. From

condition 2 in Definition 6.2.4, one has $\alpha(\|h(x) - \hat{h}(\hat{x})\|) \leq \tilde{V}(x, \hat{x}) \leq \varpi$, which readily results in $\|h(x) - \hat{h}(\hat{x})\| \leq \alpha^{-1}(\varpi) = \varepsilon$. Finally, we show the third condition of R. Consider any pair $(x, \hat{x}) \in R$, i.e., $\tilde{V}(x, \hat{x}) \leq \varpi$. From 3-a) in Definition 6.2.4, one has $\forall u$, $\forall x_d \in f(x, u)$, $\exists \hat{u}$, $\exists \hat{x}_d \in \hat{f}(\hat{x}, \hat{u})$ such that $\tilde{V}(x_d, \hat{x}_d) \leq \varpi$. It immediately follows that $(x_d, \hat{x}_d) \in R$ which satisfies condition 3-a) in Definition 4.3.1. Condition 3-b) can be proved in a similar way, which concludes the proof. The proofs for the other two relations follow the same reasoning and are omitted here. □

Now we provide the main result of this section which shows the usefulness of above-defined opacity-preserving simulation functions in terms of preserving approximate opacity across related interconnected systems.

Proposition 6.2.8. *Consider two interconnected dt-CS $\Sigma = (X, X_0, X_S, U, f, Y, h)$ and $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{f}, \hat{Y}, \hat{h})$, where $\hat{Y} \subseteq Y$, and let $\varepsilon, \delta \in \mathbb{R}_{\geq 0}$ where $\varepsilon \leq \frac{\delta}{2}$. If Σ and $\hat{\Sigma}$ admit an opacity-preserving simulation function as in Definition 6.2.4 (resp. Definition 6.2.5 or Definition 6.2.6) associated with function $\alpha \in \mathcal{K}_\infty$ and constant ϖ, then the following implication holds*

$$\hat{\Sigma} \text{ is } (\delta - 2\varepsilon)\text{-approximate opaque} \Rightarrow \Sigma \text{ is } \delta\text{-approximate opaque},$$

where $\varepsilon = \alpha^{-1}(\varpi)$.

This proposition can be proved easily by combining the results of Proposition 6.2.7 and those of Theorem 4.3.2, (resp. Theorems 4.3.7 and 4.3.9). Note that the above implication across two related systems holds for all of the three types of approximate opacity in Definition 3.3.1. This result provides us a sufficient condition for verifying approximate opacity using abstraction-based techniques.

6.2.2.2 Compositional Construction of Opacity-Preserving Simulation Functions

In the previous section, we proposed new notions of opacity-preserving simulation functions for interconnected systems using which one can check opacity using their finite abstractions. However, it is known that the construction of finite abstractions and the corresponding simulation functions for large-scale systems generally suffers from the curse of dimensionality. Motivated by this, we present here a compositional approach to establish local simulation functions for interconnected systems by composing those of the subsystems, defined below.

Definition 6.2.9. *Consider subsystems $\Sigma_i = (X_i, X_{0_i}, X_{S_i}, U_i, W_i, f_i, Y_i, h_i)$ and $\hat{\Sigma}_i = (\hat{X}_i, \hat{X}_{0_i}, \hat{X}_{S_i}, \hat{U}_i, \hat{W}_i, \hat{f}_i, \hat{Y}_i, \hat{h}_i)$ where $\hat{W}_i \subseteq W_i$ and $\hat{Y}_i \subseteq Y_i$. For $\varpi_i \in \mathbb{R}_{\geq 0}$, function $V_i : X_i \times \hat{X}_i \to \mathbb{R}_{\geq 0}$ is called a local ϖ_i-InitSOPSF from Σ_i to $\hat{\Sigma}_i$, if there exist a constant $\vartheta_i \in \mathbb{R}_{\geq 0}$, and a function $\alpha_i \in \mathcal{K}_\infty$ such that*

1. *a) $\forall x_0 \in X_{0_i} \cap X_{S_i}$, $\exists \hat{x}_{0_i} \in \hat{X}_{0_i} \cap \hat{X}_{S_i}$, s.t. $V_i(x_{0_i}, \hat{x}_{0_i}) \leq \varpi_i$;*
   b) $\forall \hat{x}_0 \in \hat{X}_{0_i} \setminus \hat{X}_{S_i}$, $\exists x_{0_i} \in X_{0_i} \setminus X_{S_i}$, s.t. $V_i(x_{0_i}, \hat{x}_{0_i}) \leq \varpi_i$;
2. *$\forall x_i \in X_i, \forall \hat{x}_i \in \hat{X}_i$, $\alpha_i(\|h_i(x_i) - \hat{h}_i(\hat{x}_i)\|) \leq V_i(x_i, \hat{x}_i)$;*

3. $\forall x_i \in X_i, \forall \hat{x}_i \in \hat{X}_i$ *s.t.* $V_i(x_i, \hat{x}_i) \leq \varpi_i$, $\forall w_i \in W_i$, $\forall \hat{w}_i \in \hat{W}_i$ *s.t.* $\|w_i - \hat{w}_i\| \leq \vartheta_i$, *the following hold:*
   a) $\forall u_i \in U_i$, $\forall x_{d_i} \in f_i(x_i, u_i, w_i)$, $\exists \hat{u}_i \in \hat{U}_i$, $\exists \hat{x}_{d_i} \in \hat{f}_i(\hat{x}_i, \hat{u}_i, \hat{w}_i)$, *s.t.* $V_i(x_{d_i}, \hat{x}_{d_i}) \leq \varpi_i$;
   b) $\forall \hat{u}_i \in \hat{U}_i$, $\forall \hat{x}_{d_i} \in \hat{f}_i(\hat{x}_i, \hat{u}_i, \hat{w}_i)$, $\exists u_i \in U_i$, $\exists x_{d_i} \in f_i(x_i, u_i, w_i)$, *s.t.* $V_i(x_{d_i}, \hat{x}_{d_i}) \leq \varpi_i$.

Similarly, we introduce new notions of local ϖ_i-CurSOPSFs and local ϖ_i-InfSOPSFs for subsystems.

Definition 6.2.10. *Consider subsystems* $\Sigma_i = (X_i, X_{0_i}, X_{S_i}, U_i, W_i, f_i, Y_i, h_i)$ *and* $\hat{\Sigma}_i = (\hat{X}_i, \hat{X}_{0_i}, \hat{X}_{S_i}, \hat{U}_i, \hat{W}_i, \hat{f}_i, \hat{Y}_i, \hat{h}_i)$ *where* $\hat{W}_i \subseteq W_i$ *and* $\hat{Y}_i \subseteq Y_i$. *For* $\varpi_i \in \mathbb{R}_{\geq 0}$, *function* $V_i : X_i \times \hat{X}_i \to \mathbb{R}_{\geq 0}$ *is called a local* ϖ_i*-CurSOPSF from* Σ_i *to* $\hat{\Sigma}_i$, *if there exist a constant* $\vartheta_i \in \mathbb{R}_{\geq 0}$, *and a function* $\alpha_i \in \mathcal{K}_\infty$ *such that*

1. $\forall x_{0_i} \in X_{0_i}$, $\exists \hat{x}_{0_i} \in \hat{X}_{0_i}$, *s.t.* $V_i(x_{0_i}, \hat{x}_{0_i}) \leq \varpi_i$;
2. $\forall x_i \in X_i, \forall \hat{x}_i \in \hat{X}_i$, $\alpha_i(\|h_i(x_i) - \hat{h}_i(\hat{x}_i)\|) \leq V_i(x_i, \hat{x}_i)$;
3. $\forall x_i \in X_i, \forall \hat{x}_i \in \hat{X}_i$ *s.t.* $V_i(x_i, \hat{x}_i) \leq \varpi_i$, $\forall w_i \in W_i$, $\forall \hat{w}_i \in \hat{W}_i$ *s.t.* $\|w_i - \hat{w}_i\| \leq \vartheta_i$, *the following hold:*
   a) $\forall u_i \in U_i$, $\forall x_{d_i} \in f_i(x_i, u_i, w_i)$, $\exists \hat{u}_i \in \hat{U}_i$, $\exists \hat{x}_{d_i} \in \hat{f}_i(\hat{x}_i, \hat{u}_i, \hat{w}_i)$, *s.t.* $V_i(x_{d_i}, \hat{x}_{d_i}) \leq \varpi_i$;
   b) $\forall u_i \in U_i$, $\forall x_{d_i} \in f_i(x_i, u_i, w_i)$ *s.t.* $x_{d_i} \in X_{S_i}$, $\exists \hat{u}_i \in \hat{U}_i$, $\exists \hat{x}_{d_i} \in \hat{f}_i(\hat{x}_i, \hat{u}_i, \hat{w}_i)$ *with* $\hat{x}_{d_i} \in \hat{X}_{S_i}$, *s.t.* $V_i(x_{d_i}, \hat{x}_{d_i}) \leq \varpi_i$;
   c) $\forall \hat{u}_i \in \hat{U}_i$, $\forall \hat{x}_{d_i} \in \hat{f}_i(\hat{x}_i, \hat{u}_i, \hat{w}_i)$, $\exists u_i \in U_i$, $\exists x_{d_i} \in f_i(x_i, u_i, w_i)$, *s.t.* $V_i(x_{d_i}, \hat{x}_{d_i}) \leq \varpi_i$;
   d) $\forall \hat{u}_i \in \hat{U}_i$, $\forall \hat{x}_{d_i} \in \hat{f}_i(\hat{x}_i, \hat{u}_i, \hat{w}_i)$ *s.t.* $\hat{x}_{d_i} \in \hat{X}_i \setminus \hat{X}_{S_i}$, $\exists u_i \in U_i$, $\exists x_{d_i} \in f_i(x_i, u_i, w_i)$ *with* $x_{d_i} \in X_i \setminus X_{S_i}$, *s.t.* $V_i(x_{d_i}, \hat{x}_{d_i}) \leq \varpi_i$.

Definition 6.2.11. *Consider subsystems* $\Sigma_i = (X_i, X_{0_i}, X_{S_i}, U_i, W_i, f_i, Y_i, h_i)$ *and* $\hat{\Sigma}_i = (\hat{X}_i, \hat{X}_{0_i}, \hat{X}_{S_i}, \hat{U}_i, \hat{W}_i, \hat{f}_i, \hat{Y}_i, \hat{h}_i)$ *where* $\hat{W}_i \subseteq W_i$ *and* $\hat{Y}_i \subseteq Y_i$. *For* $\varpi_i \in \mathbb{R}_{\geq 0}$, *a function* $V_i : X_i \times \hat{X}_i \to \mathbb{R}_{\geq 0}$ *is called a local* ϖ_i*-InfSOPSF from* Σ_i *to* $\hat{\Sigma}_i$, *if it is both a local* ϖ_i*-InitSOPSF and a local* ϖ_i*-CurSOPSF from* Σ_i *to* $\hat{\Sigma}_i$.

If there exists a local opacity-preserving simulation function from Σ_i to $\hat{\Sigma}_i$, and $\hat{\Sigma}_i$ is finite, $\hat{\Sigma}_i$ is called a local finite abstraction of the concrete subsystem Σ_i. Note that the local simulation functions are mainly proposed for constructing overall simulation functions for networks and are not directly used for deducing the preservation of approximate opacity between subsystems. Next, we show how to compose the above-defined local opacity-preserving simulation functions so that they can be used to quantify the distance between two networks.

Theorem 6.2.12. Consider an interconnected dt-CS $\Sigma = \mathcal{I}(\Sigma_1, \ldots, \Sigma_N)$ induced by $N \in \mathbb{N}_{\geq 1}$ subsystems Σ_i. Assume that each Σ_i and its abstraction $\hat{\Sigma}_i$ admit a local ϖ_i-InitSOPSF (resp. ϖ_i-CurSOPSF or ϖ_i-InfSOPSF) V_i. Let $\varpi = \max_i \varpi_i$. If

$$\alpha_j^{-1}(\varpi_j) + \phi_{ij} \leq \vartheta_i, \forall i \in [1; N], \forall j \neq i, \tag{6.2.7}$$

where ϕ_{ij} is an internal input quantization parameter for constructing the finite abstractions $\hat{\Sigma}_i$, then, function

$$\tilde{V}(x, \hat{x}) := \max_i \{\frac{\varpi}{\varpi_i} V_i(x_i, \hat{x}_i)\}, \tag{6.2.8}$$

is an ϖ-InitSOPSF (resp. ϖ-CurSOPSF or ϖ-InfSOPSF) from Σ to $\hat{\Sigma} = \hat{\mathcal{I}}(\hat{\Sigma}_1, \ldots, \hat{\Sigma}_N)$.

Proof. First, we show that condition 1a) in Definition 6.2.4 holds. Consider any $x_0 = [x_{0_1}; \ldots; x_{0_N}] \in X_0 \cap X_S$. For any subsystem Σ_i and the corresponding abstraction $\hat{\Sigma}_i$, from the definition of local ϖ_i-InitSOPSF V_i, we have $\forall x_{0_i} \in X_{0_i} \cap X_{S_i}$, $\exists \hat{x}_{0_i} \in \hat{X}_{0_i} \cap \hat{X}_{S_i}$: $V_i(x_{0_i}, \hat{x}_{0_i}) \leq \varpi_i$. Then, from the definition of $\tilde{V}$ as in (6.2.8) we get $\tilde{V}(x_0, \hat{x}_0) \leq \varpi$, where $\hat{x}_0 = [\hat{x}_{0_1}; \ldots; \hat{x}_{0_N}] \in \hat{X}_0 \cap \hat{X}_S$. Thus, condition 1a) in Definition 6.2.4 holds. Condition 1b) can be proved in the same way thus is omitted here. Now, we show that condition 2 in Definition 6.2.4 holds for some $\mathcal{K}_\infty$ function α. Consider any $x = [x_1; \ldots; x_N] \in X$ and $\hat{x} = [\hat{x}_1; \ldots; \hat{x}_N] \in \hat{X}$. Then, using condition 2 in Definition 6.2.9, one gets

$$\begin{aligned}\|h(x) - \hat{h}(\hat{x})\| &= \max_i \{\|h_{ii}(x_i) - \hat{h}_{ii}(\hat{x}_i)\|\} \leq \max_i \{\|h_i(x_i) - \hat{h}_i(\hat{x}_i)\|\} \\ &\leq \max_i \{\alpha_i^{-1}(V_i(x_i, \hat{x}_i))\} \leq \hat{\alpha}(\max_i \{\frac{\varpi}{\varpi_i} V_i(x_i, \hat{x}_i)\}),\end{aligned}$$

where $\hat{\alpha}(s) = \max_i \{\alpha_i^{-1}(s)\}, \forall s \in \mathbb{R}_{\geq 0}$. By defining $\alpha = \hat{\alpha}^{-1}$, one obtains

$$\alpha(\|h(x) - \hat{h}(\hat{x})\|) \leq \tilde{V}(x, \hat{x}),$$

which satisfies condition 2. Next, we show that condition 3 holds. Let us consider any $x = [x_1; \ldots; x_N] \in X$ and $\hat{x} = [\hat{x}_1; \ldots; \hat{x}_N] \in \hat{X}$ such that $\tilde{V}(x, \hat{x}) \leq \varpi$. It can be seen that from the construction of $\tilde{V}$ in (6.2.8), we get $V_i(x_i, \hat{x}_i) \leq \varpi_i$ holds, $\forall i \in [1; N]$. For each pair of subsystems Σ_i and $\hat{\Sigma}_i$, the internal inputs satisfy the chain of inequality:

$$\begin{aligned}\|w_i - \hat{w}_i\| =& \max_{j \neq i} \{\|w_{ij} - \hat{w}_{ij}\|\} = \max_{j \neq i} \{\|y_{ji} - \hat{y}_{ji} + \hat{y}_{ji} - \hat{w}_{ij}\|\} \\ \leq& \max_{j \neq i} \{\|y_{ji} - \hat{y}_{ji}\| + \phi_{ij}\} \leq \max_{j \neq i} \{\|h_j(x_j) - \hat{h}_j(\hat{x}_j)\| + \phi_{ij}\} \\ \leq& \max_{j \neq i} \{\alpha_j^{-1}(V_j(x_j, \hat{x}_j)) + \phi_{ij}\} \leq \max_{j \neq i} \{\alpha_j^{-1}(\varpi_j) + \phi_{ij}\}.\end{aligned}$$

Using (6.2.7), one has $\|w_i - \hat{w}_i\| \leq \vartheta_i$. Therefore, by Definition 6.2.9 for each pair of subsystems Σ_i and $\hat{\Sigma}_i$, one has $\forall u_i \in U_i$ $\forall x_{d_i} \in f_i(x_i, u_i, w_i)$, there exists $\hat{u}_i \in \hat{U}_i$ and $\hat{x}_{d_i} \in \hat{f}_i(\hat{x}_i, \hat{u}_i, \hat{w}_i)$ such that $V_i(x_{d_i}, \hat{x}_{d_i}) \leq \varpi_i$. As a result, we get $\forall u = [u_1; \ldots; u_N] \in U$, $\forall x_d \in f(x, u)$, there exists $\hat{u} = [\hat{u}_1; \ldots; \hat{u}_N] \in \hat{U}$ and $\hat{x}_d \in \hat{f}(\hat{x}, \hat{u})$ such that $\tilde{V}(x_d, \hat{x}_d) := \max_i\{\frac{\varpi}{\varpi_i} V_i(x_{d_i}, \hat{x}_{d_i})\} \leq \varpi$. Therefore, condition 3a) in Definition 6.2.4 is satisfied with $\varpi = \max_i \varpi_i$. The proof of condition 3b) uses the same reasoning as that of 3a) and is omitted here. Therefore, we conclude that $\tilde{V}$ is an ϖ-InitSOPSF from Σ to $\hat{\Sigma}$. In a similar way, one can prove that $\tilde{V}$ is also an ϖ-CurSOPSF (resp. ϖ-InfSOPSF) from Σ to $\hat{\Sigma}$. □

In the sequel, we will impose conditions on the dynamics of the subsystems such that one can establish proper finite abstractions together with their corresponding local opacity-preserving simulation functions for all of the subsystems.

6.2.3 Compositionality Results

In this section, we present a method to construct local finite abstractions, together with the corresponding local opacity-preserving simulation functions for the concrete subsystems satisfying certain stability property. We consider each subsystem $\Sigma_i = (X_i, X_{0_i}, X_{S_i}, U_i, W_i, f_i, Y_i, h_i)$ as an infinite, deterministic discrete-time control subsystem with $X_{0_i} = X_i$. We assume the output map h_i of Σ_i satisfies the following general Lipschitz assumption $\|h_i(x_i) - h(x_i')\| \leq \ell(\|x_i - x_i'\|)$, for all $x_i, x_i' \in X_i$, where $\ell \in \mathcal{K}_\infty$.

6.2.3.1 Construction of Local Finite Abstractions

Note that throughout this subsection, we will work on subsystems rather than the overall network. However, we omit index i of subsystems throughout the text for the sake of better readability, e.g., we write Σ instead of Σ_i. The opacity-preserving simulation functions between Σ and its local finite abstraction is established under the assumption that Σ is incrementally input-to-state stable (δ-ISS) [10] as defined next. Note that the definition of incremental input-to-state stability for a subsystem is slightly different from that in Definition 2.4.3 by incorporating the influence of internal inputs.

Definition 6.2.13. *A discrete-time control subsystem $\Sigma = (X, X_0, X_S, U, W, f, Y, h)$ is δ-ISS if there exist functions $\mathcal{G} : X \times X \to \mathbb{R}_{\geq 0}$, $\underline{\alpha}$, $\overline{\alpha}$, κ, ρ_{int}, $\rho_{ext} \in \mathcal{K}_\infty$, such that $\forall x, x' \in X$, $\forall u, u' \in U$, $\forall w, w' \in W$,*

$$\underline{\alpha}(\|x - x'\|) \leq \mathcal{G}(x, x') \leq \overline{\alpha}(\|x - x'\|), \tag{6.2.9}$$

$$\begin{aligned}&\mathcal{G}(f(x, u, w), f(x', u', w')) - \mathcal{G}(x, x')\\ &\leq -\kappa(\mathcal{G}(x, x')) + \rho_{int}(\|w - w'\|) + \rho_{ext}(\|u - u'\|).\end{aligned} \tag{6.2.10}$$

We additionally assume that there exists a function $\hat{\gamma} \in \mathcal{K}_\infty$ such that $\forall x, x', x'' \in X$,

$$\mathcal{G}(x, x') \leq \mathcal{G}(x, x'') + \hat{\gamma}(\|x' - x''\|), \tag{6.2.11}$$

for $\mathcal{G}$ defined in Definition 6.2.13. Note that in most real applications, the state set X is a compact subset of $\mathbb{R}^n$ and, hence, condition (6.2.11) is not restrictive.

Now, we construct a local finite abstraction of a δ-ISS discrete-time control subsystem $\Sigma = (X, X_0, X_S, U, W, f, Y, h)$. For the remainder of the chapter, we assume that sets X, X_0, X_S, W, and U are of the form of finite unions of boxes. Consider a tuple $q = (\eta, \theta, \mu, \phi)$ of parameters, where $0 \leq \eta \leq \min\{span(X_S), span(X \setminus X_S)\}$ is the state set quantization, $0 \leq \mu < span(U)$ is the external input set quantization, ϕ is the internal input set quantization parameter, where $0 \leq \|\phi\| \leq span(W)$, and $\theta \in \mathbb{R}_{\geq 0}$ is a design parameter. A local finite abstraction can be represented as the tuple $\hat{\Sigma} = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{W}, \hat{f}, \hat{Y}, \hat{h})$, where $\hat{X} = \hat{X}_0 = [X]_\eta$, $\hat{X}_S = [X_S^\theta]_\eta$, $\hat{U} = [U]_\mu$, $\hat{W} = [W]_\phi$, $\hat{Y} = \{h(\hat{x}) | \hat{x} \in \hat{X}\}$, $\hat{h}(\hat{x}) = h(\hat{x})$, $\forall \hat{x} \in \hat{X}$, and $\hat{x}_d \in \hat{f}(\hat{x}, \hat{u}, \hat{w})$ if and only if $\|\hat{x}_d - f(\hat{x}, \hat{u}, \hat{w})\| \leq \eta$, where $X_S^\theta = \{x \in X \mid \exists \bar{x} \in X_S, \|x - \bar{x}\| \leq \theta\}$ denotes the θ-expansion of X_S.

Next, we show that if the abstraction $\hat{\Sigma}$ of a δ-ISS Σ is constructed with the tuple of parameters satisfying some conditions, then function $\mathcal{G}$ in Definition 6.2.13 is a local InitSOPSF (resp. CurSOPSF or InfSOPSF) from Σ to $\hat{\Sigma}$.

Theorem 6.2.14. Consider a δ-ISS discrete-time control subsystem $\Sigma = (X, X_0, X_S, U, W, f, Y, h)$ as in Definition 6.2.1 with function $\mathcal{G}$ satisfying (6.2.9)-(6.2.11) with $\mathcal{K}_\infty$ functions $\underline{\alpha}, \overline{\alpha}, \kappa, \rho_{int}, \rho_{ext}, \hat{\gamma}$. For any design parameters $\varpi, \vartheta \in \mathbb{R}_{\geq 0}$, let $\hat{\Sigma}$ be a finite abstraction of Σ with a tuple $q = (\eta, 0, \mu, \phi)$ of parameters satisfying

$$\eta \leq \min\{\hat{\gamma}^{-1}[\kappa(\varpi) - \rho_{int}(\vartheta) - \rho_{ext}(\mu)], \overline{\alpha}^{-1}(\varpi)\}. \tag{6.2.12}$$

Then, $\mathcal{G}$ is a local ϖ-InitSOPSF from Σ to $\hat{\Sigma}$ and from $\hat{\Sigma}$ to Σ.

Proof. We start by proving condition 1 in Definition 6.2.9. Consider any initial and secret state $x_0 \in X_0 \cap X_S$ in Σ. Since $\eta \leq span(X_S)$, $X_S \subseteq \bigcup_{p \in [X_S]_\eta} \mathcal{B}_\eta(p)$, then for every $x \in X_S$ there always exists $\hat{x} \in \hat{X}_S$ such that $\|x - \hat{x}\| \leq \eta$. Hence, there exists $\hat{x}_0 \in \hat{X}_0 \cap \hat{X}_S$ with $\mathcal{G}(x_0, \hat{x}_0) \leq \overline{\alpha}(\|x_0 - \hat{x}_0\|) \leq \overline{\alpha}(\eta)$ by (6.2.9), and condition 1(a) in Definition 6.2.9 is satisfied with $\varpi \geq \overline{\alpha}(\eta)$ by (6.2.12). For every $\hat{x}_0 \in \hat{X}_0 \setminus \hat{X}_S$, by choosing $x_0 = \hat{x}_0$ which is also inside $X_0 \setminus X_S$, we get $\mathcal{G}(x_0, \hat{x}_0) = 0 \leq \varpi$. Hence, condition 1(b) in Definition 6.2.9 holds as well. Next, we show that condition 2 in Definition 6.2.9 holds. Since Σ is incrementally input-to-state stable as in (6.2.9), and given the Lipschitz assumption on h, $\forall x \in X$ and $\forall \hat{x} \in \hat{X}$, we have

$$\|h(x) - \hat{h}(\hat{x})\| \leq \ell(\|x - \hat{x}\|) \leq \ell \circ \underline{\alpha}^{-1}(\mathcal{G}(x, \hat{x})).$$

Let us define $\alpha = (\ell \circ \underline{\alpha}^{-1})^{-1}$. Then one obtains that condition 2 in Definition 6.2.9 is satisfied with

$$\alpha(\|h(x) - \hat{h}(\hat{x})\|) \leq \mathcal{G}(x, \hat{x}).$$

Now we show condition 3 in Definition 6.2.9. From (6.2.11), $\forall x \in X, \forall \hat{x} \in \hat{X}, \forall u \in U, \forall \hat{u} \in \hat{U}, \forall w \in W, \forall \hat{w} \in \hat{W}$, we have for any $\hat{x}_d \in \hat{f}(\hat{x}, \hat{u}, \hat{w})$

$$\mathcal{G}(x_d, \hat{x}_d)) \leq \mathcal{G}(x_d, f(\hat{x}, \hat{u}, \hat{w})) + \hat{\gamma}(\|\hat{x}_d - f(\hat{x}, \hat{u}, \hat{w})\|),$$

where[1] $x_d = f(x, u, w)$. From the structure of abstraction, the above inequality reduces to

$$\mathcal{G}(x_d, \hat{x}_d) \leq \mathcal{G}(x_d, f(\hat{x}, \hat{u}, \hat{w})) + \hat{\gamma}(\eta).$$

Note that by (6.2.10), we get

$$\mathcal{G}(x_d, f(\hat{x}, \hat{u}, \hat{w})) - \mathcal{G}(x, \hat{x}) \leq -\kappa(\mathcal{G}(x, \hat{x})) + \rho_{ext}(\|u - \hat{u}\|) + \rho_{int}(\|w - \hat{w}\|).$$

Hence, $\forall x \in X, \forall \hat{x} \in \hat{X}, \forall u \in U, \forall \hat{u} \in \hat{U}$, $\forall w \in W, \forall \hat{w} \in \hat{W}$, one obtains

$$\mathcal{G}(x_d, \hat{x}_d) - \mathcal{G}(x, \hat{x}) \leq -\kappa(\mathcal{G}(x, \hat{x})) + \rho_{ext}(\|u - \hat{u}\|) + \rho_{int}(\|w - \hat{w}\|) + \hat{\gamma}(\eta), \quad (6.2.13)$$

for any $\hat{x}_d \in \hat{f}(\hat{x}, \hat{u}, \hat{w})$. Now, we show condition 3(a) in Definition 6.2.9. Let us consider any $x \in X$ and any $\hat{x} \in \hat{X}$ satisfying $\mathcal{G}(x, \hat{x}) \leq \varpi$, and any $w \in W$ and $\hat{w}$ such that $\|\hat{w} - w\| \leq \vartheta$. By the structure of $\hat{U} = [U]_\mu$, for any $u \in U$, there always exists $\hat{u}$ satisfying $\|\hat{u} - u\| \leq \mu$. By combining (6.2.13) with (6.2.12), for any $x_d = f(x, u, w)$ and any $\hat{x}_d \in \hat{f}(\hat{x}, \hat{u}, \hat{w})$, the following inequality holds:

$$\mathcal{G}(x_d, \hat{x}_d) \leq (\mathrm{id} - \kappa)(\varpi) + \rho_{ext}(\mu) + \rho_{int}(\vartheta) + \hat{\gamma}(\eta) \leq \varpi. \quad (6.2.14)$$

Hence, condition 3(a) is satisfied. Similarly, for any $\hat{u}$, by choosing $u = \hat{u}$, for any $\hat{x}_d \in \hat{f}(\hat{x}, \hat{u}, \hat{w})$, condition 3(b) in Definition 6.2.10 is also satisfied with $\mathcal{G}(x_d, \hat{x}_d) \leq (\mathrm{id} - \kappa)(\varpi) + \rho_{int}(\vartheta) + \hat{\gamma}(\eta) \leq \varpi$, where $x_d = f(x, u, w)$. Therefore, we conclude that $\mathcal{G}$ is a ϖ-InitSOPSF from Σ to $\hat{\Sigma}$. □

Next, we provide a similar result as in Theorem 6.2.14, but tailored to current-state and infinite-step opacity.

Theorem 6.2.15. Consider a δ-ISS discrete-time control subsystem $\Sigma = (X, X_0, X_S, U, W, f, Y, h)$ as in Definition 6.2.1 with function $\mathcal{G}$ satisfying (6.2.9)-(6.2.11) with $\mathcal{K}_\infty$ functions $\underline{\alpha}, \overline{\alpha}, \kappa, \rho_{int}, \rho_{ext}, \hat{\gamma}$. For any design parameters $\varpi, \vartheta \in \mathbb{R}_{\geq 0}$, let $\hat{\Sigma}$ be a finite abstraction of Σ with a tuple $q = (\eta, \theta, \mu, \phi)$ of parameters satisfying

$$\eta \leq \min\{\hat{\gamma}^{-1}[\kappa(\varpi) - \rho_{int}(\vartheta) - \rho_{ext}(\mu)], \overline{\alpha}^{-1}(\varpi)\}; \quad (6.2.15)$$

$$\underline{\alpha}^{-1}(\varpi) \leq \theta. \quad (6.2.16)$$

Then, $\mathcal{G}$ is a local ϖ-CurSOPSF (resp. InfSOPSF) from Σ to $\hat{\Sigma}$.

[1] In this section, we assume that Σ is deterministic.

Proof. We start by proving condition 1 in Definition 6.2.10. Since $\hat{X} = \hat{X}_0 = [X]_\eta = [X_0]_\eta$, $X_0 \subseteq \bigcup_{p\in\hat{X}_0} \mathcal{B}_\eta(p)$, then for every initial state $x_0 \in X_0$ in Σ there always exists $\hat{x}_0 \in \hat{X}_0$ in $\widehat{\Sigma}$ such that $\|\hat{x}_0 - x_0\| \leq \eta$. Hence, one gets $\mathcal{G}(x_0, \hat{x}_0) \leq \overline{\alpha}(\|x_0 - \hat{x}_0\|) \leq \overline{\alpha}(\eta)$ by (6.2.9), and by using (6.2.15) condition 1 in Definition 6.2.10 is satisfied with $\varpi \geq \overline{\alpha}(\eta)$. The proof for conditions 2, 3(a), and 3(c) in Definition 6.2.10 is similar to that of Theorem 6.2.14, and is omitted here. For condition 3(b), let us consider any $u \in U$ s.t. $x_d = f(x, u, w) \in X_S$. Again, by choosing any $\hat{u}$ satisfying $\|\hat{u} - u\| \leq \mu$, we obtain $\mathcal{G}(x_d, \hat{x}_d) \leq \varpi$. Additionally, by (6.2.9) one gets

$$\|x_d - \hat{x}_d\| \leq \underline{\alpha}^{-1}(\mathcal{G}(x_d, \hat{x}_d)) \leq \underline{\alpha}^{-1}(\varpi). \tag{6.2.17}$$

As one can see from the structure of the abstraction, where $\hat{X}_S = [X_S^\theta]_\eta$ and using $\theta \geq \underline{\alpha}^{-1}(\varpi)$ in (6.2.16), from $x_d \in X_S$ one concludes that $\hat{x}_d \in \hat{X}_S$, which shows that condition 3(b) holds as well. Condition 3(d) can be proved similarly, which shows that $\mathcal{G}$ is a ϖ-approximate current-state opacity-preserving simulation function from Σ to $\hat{\Sigma}$. The proof for ϖ-approximate InfSOPSF follows the same reasoning as those in Theorems 6.2.14 and 6.2.15. □

Remark 6.2.16. *Note that the proposed local simulation functions provide one-sided relations since condition 1 in Definition 6.2.9 (or 6.2.10) is not symmetric. On the other hand, the two-sided (symmetric) decay condition 3 in Definition 6.2.9 (or 6.2.10) is similar to the* approximate bisimulation relation *proposed in [52]. We refer interested readers to [222, Examples 3.5 and 3.6], where the two-sided conditions are shown to be necessary to ensure the preservation of opacity. Therefore, in order to find suitable local opacity-preserving simulation functions, the δ-ISS assumption is still required for the subsystems. Notice that under the δ-ISS assumption, we showed that concrete system and its abstraction simulates each other in terms of preserving initial-state opacity (cf. Theorem 6.2.14). However, in the case of CurSOPSF and InfSOPSF, having δ-ISS property only ensures that the abstract system simulates the concrete one and not the other direction (cf. Theorem 6.2.15).*

One can observe that in order to satisfy conditions (6.2.7) and (6.2.12) (resp. (6.2.15)) simultaneously, the interconnected system must hold some property. In the next subsection, we will discuss about the inherent property that the interconnected system should hold such that one can design suitable quantization parameters to satisfy these competing conditions at the same time.

6.2.3.2 Compositional Construction of Opacity-Preserving Finite Abstractions

In this subsection, we exploit the interconnection topology of the overall network and employ the knowledge from graph theory as an essential tool in our main result. Here, we first introduce some terminologies that will be used later based on the notion of strongly connected components (SCCs) [13], which are used to represent sub-networks of an interconnected system.

Consider an interconnected dt-CS $\Sigma = \mathcal{I}(\Sigma_1, \ldots, \Sigma_N)$ induced by $N \in \mathbb{N}_{\geq 1}$ δ-ISS subsystems Σ_i. We denote by $G = (I, E)$ the directed graph associated with Σ, where $I = [1; N]$ is the set of vertices with each vertex $i \in I$ labelled with subsystem Σ_i, and $E \subseteq I \times I$ is the set of ordered pairs (i, j), $\forall i, j \in I$, with $y_{ji} \neq 0$. The SCCs of G are denoted by $\bar{G}_k = (I_k, E_k)$, $k \in [1; \bar{N}]$, where $\bar{N}$ is the number of SCCs in G. For any $\bar{G}_k$, we set $I_k = \{k_1, \ldots, k_{\bar{N}_k}\}$ and $\bar{N}_k = \text{card}(I_k)$. We denote by $\mathcal{N}_I(i) = \{j \in I | \exists (i, j) \in E\}$ and $\mathcal{M}_I(i) = \{j \in I | \exists (j, i) \in E\}$ the set of vertices in I which are direct predecessors of i and direct successors of i, respectively. We denote by $\text{BSCC}(G)$ the collection of bottom SCCs of G from which no vertex in G outside $\bar{G}_k$ is reachable.

Now, we raise the following small-gain type assumption which is essential for the main compositionality result.

Assumption 6.2.17. *Consider an interconnected dt-CS $\Sigma = \mathcal{I}(\Sigma_1, \ldots, \Sigma_N)$ induced by $N \in \mathbb{N}_{\geq 1}$ δ-ISS subsystems Σ_i which is associated with a directed graph G. Assume that each Σ_i and its abstraction $\hat{\Sigma}_i$ admit a local ϖ_i-InitSOPSF (resp. CurSOPSF or InfSOPSF) $\mathcal{G}_i$, together with functions κ_i, α_i, and ρ_{inti} as appeared in Definition 6.2.9 (resp. Definition 6.2.10 or Definition 6.2.11) and Definition 6.2.13. For every SCC $\bar{G}_k$ in G, we define*

$$\gamma_{ij} = \begin{cases} \kappa_i^{-1} \circ \rho_{inti} \circ \alpha_j^{-1} & \text{if } j \in \mathcal{N}_{I_k}(i), \\ 0 & \text{otherwise}, \end{cases} \tag{6.2.18}$$

where $\mathcal{N}_{I_k}(i) = \{j \in I_k | \exists (i, j) \in E\}$, $\forall i, j \in I_k$. We assume that for every $\bar{G}_k$, $k \in [1; \bar{N}]$, the following holds

$$\gamma_{i_1 i_2} \circ \gamma_{i_2 i_3} \circ \cdots \circ \gamma_{i_{r-1} i_r} \circ \gamma_{i_r i_1} < \text{id}, \tag{6.2.19}$$

$\forall (i_1, \ldots, i_r) \in \{k_1, \ldots, k_{\bar{N}_k}\}^r$, where $r \in \{1, \ldots, \bar{N}_k\}$.

Now, we provide the next main result showing that under the above assumption, one can always compositionally design local quantization parameters such that conditions (6.2.7) and (6.2.12) (resp. (6.2.15)) are fulfilled simultaneously.

Theorem 6.2.18. *Suppose that Assumption 6.2.17 holds. Then, for any desired precision $\varpi \in \mathbb{R}_{>0}$ as in Definition 6.2.4 (resp. Definition 6.2.5 or 6.2.6), there always exist quantization parameters η_i, μ_i, ϕ_i, $\forall i \in [1; N]$, such that* (6.2.7) *and* (6.2.12) *(resp.* (6.2.15)*) are satisfied simultaneously, where the local parameters $\vartheta_i \in \mathbb{R}_{>0}$ and $\varpi_i \in \mathbb{R}_{>0}$, $\forall i \in [1; N]$, are obtained from Algorithm 1.*

Proof. First, let us note that the small-gain type condition (6.2.19) implies that for each $\bar{G}_k$, there exists $\sigma_i \in \mathcal{K}_\infty$ satisfying, $\forall i \in I_k$,

$$\max_{j \in \mathcal{N}_{I_k}(i)} \{\gamma_{ij} \circ \sigma_j\} < \sigma_i; \tag{6.2.20}$$

see [39, Theorem 5.2]. Now, given a desired precision ϖ, we apply Algorithm 1 to design the pair of parameters (ϖ_i, ϑ_i) for all of the subsystems. In order to show that

Algorithm 1: Compositional design of local parameters $\varpi_i \in \mathbb{R}_{>0}$ and $\vartheta_i \in \mathbb{R}_{>0}$, $\forall i \in [1; N]$

Input: The desired precision $\varpi \in \mathbb{R}_{>0}$; the directed graph G composed of SCCs $\bar{G}_k$ and functions σ_{k_i} $\forall i \in I_k$ satisfying (6.2.20) for $\bar{G}_k$, $\forall k \in [1; \bar{N}]$; the functions $\mathcal{G}_i$ equipped with functions κ_i, α_i, and ρ_{inti}, $\forall i \in [1; N]$.

Output: $\varpi_i \in \mathbb{R}_{>0}$ and $\vartheta_i \in \mathbb{R}_{>0}$, $\forall i \in [1; N]$.

1 Set $\varpi_i := \infty$, $\vartheta_i := \infty$, $\forall i \in [1; N]$, $\forall k \in [1; \bar{N}]$, $G^* = G$
2 **while** $G^* \neq \varnothing$ **do**
3 **foreach** $\bar{G}_k \in \mathrm{BSCC}(G^*)$ **do**
4 **if** $G^* = G$ **then**
 `/* Graph` G `represents the entire network */`
5 **if** $\bar{N}_k > 1$ **then** choose $r \in \mathbb{R}_{>0}$ s.t. $\max_{i \in I_k}\{\sigma_i(r)\} = \varpi$; set $\varpi_i = \sigma_i(r)$, choose ϕ_{ij} s.t. $\max_{j \in \mathcal{N}_{I_k}(i)}\{\phi_{ij}\} < \rho_{inti}^{-1} \circ \kappa_i(\varpi_i) - \max_{j \in \mathcal{N}_{I_k}(i)}\{\alpha_j^{-1}(\varpi_j)\}$, $\forall i, j \in I_k$, set $\vartheta_i = \max_{j \in \mathcal{N}_{I_k}(i)}\{\alpha_j^{-1}(\varpi_j) + \phi_{ij}\}$, $\forall i \in I_k$, and choose $\phi_{ij} < \vartheta_i, \forall i \in I_k, \forall j \in \mathcal{N}_{I \setminus I_k}(i)$;
6 **else** `/* The SCC contains only 1 subsystem */`
7 set $\varpi_i = \varpi$, choose $\vartheta_i \in \mathbb{R}_{>0}$ s.t. $\vartheta_i < \rho_{inti}^{-1} \circ \kappa_i(\varpi_i)$, $i \in I_k$; choose $\phi_{ij} < \vartheta_i, \forall i \in I_k, \forall j \in \mathcal{N}_{I \setminus I_k}(i)$;
8 **else**
9 **if** $\bar{N}_k > 1$ **then** choose $r \in \mathbb{R}_{>0}$ s.t. $\sigma_i(r) \leq \alpha_i(\min_{j \in \mathcal{M}_{I \setminus I_k}(i)}\{\vartheta_j - \phi_{ji}\})$, $\forall i \in I_k$ with $\mathcal{M}_{I \setminus I_k}(i) \neq \varnothing$; set $\varpi_i = \sigma_i(r)$, choose ϕ_{ij} s.t. $\max_{j \in \mathcal{N}_{I_k}(i)}\{\phi_{ij}\} < \rho_{inti}^{-1} \circ \kappa_i(\varpi_i) - \max_{j \in \mathcal{N}_{I_k}(i)}\{\alpha_j^{-1}(\varpi_j)\}, \forall i, j \in I_k$, set $\vartheta_i = \max_{j \in \mathcal{N}_{I_k}(i)}\{\alpha_j^{-1}(\varpi_j) + \phi_{ij}\}$, $\forall i \in I_k$, and choose $\phi_{ij} < \vartheta_i$, $\forall i \in I_k$, $\forall j \in \mathcal{N}_{I \setminus I_k}(i)$;
10 **else** `/* The SCC contains only 1 subsystem */`
11 set $\varpi_i \leq \alpha_i(\min_{j \in \mathcal{M}_{I \setminus I_k}(i)}\{\vartheta_j - \phi_{ji}\})$, choose $\vartheta_i \in \mathbb{R}_{>0}$ s.t. $\vartheta_i < \rho_{inti}^{-1} \circ \kappa_i(\varpi_i)$, $i \in I_k$; choose $\phi_{ij} < \vartheta_i, \forall i \in I_k$, $\forall j \in \mathcal{N}_{I \setminus I_k}(i)$;
12 **end**
13 **end**
14 $G^* = G^* \setminus \mathrm{BSCC}(G^*)$;
15 **end**

the algorithm guarantees the simultaneous satisfaction of conditions (6.2.7) and (6.2.12) (resp. (6.2.15)), let us consider different scenarios of the SCCs. First, we consider the SCCs which are composed of only 1 subsystem, i.e $\bar{N}_k = 1$. From lines 6 and 9, one observes that the selections of ϖ_i and ϑ_i for each subsystem immediately ensure that $\kappa_i(\varpi_i) - \rho_{inti}(\vartheta_i) > 0$, which implies that there always exist quantization parameters

η_i, μ_i to satisfy (6.2.12) (resp. (6.2.15)). Next, let us consider the SCCs with more than 1 subsystems, i.e $\bar{N}_k > 1$. Suppose that for each $\bar{G}_k$, we are given functions $\sigma_i \in \mathcal{K}_\infty, \forall i \in I_k$ satisfying (6.2.20). From (6.2.18) and (6.2.20), we have

$$\begin{aligned}&\max_{j\in\mathcal{N}_{I_k}(i)}\{\gamma_{ij}\circ\sigma_j\} < \sigma_i \Longrightarrow \max_{j\in\mathcal{N}_{I_k}(i)}\{\kappa_i^{-1}\circ\rho_{inti}\circ\alpha_j^{-1}\circ\sigma_j\} < \sigma_i\\ &\Longrightarrow \rho_{inti}\circ\max_{j\in\mathcal{N}_{I_k}(i)}\{\alpha_j^{-1}\circ\sigma_j\} < \kappa_i\circ\sigma_i,\end{aligned} \tag{6.2.21}$$

which holds for each $i \in I_k$. Now, let us set $\varpi_i = \sigma_i(r), \forall i \in I_k$, where r is chosen under the criteria in lines 5 and 8, and choose the internal input quantization parameters ϕ_{ij} such that $\forall i, j \in I_k$

$$\max_{j\in\mathcal{N}_{I_k}(i)}\{\phi_{ij}\} < \rho_{inti}^{-1}\circ\kappa_i(\varpi_i) - \max_{j\in\mathcal{N}_{I_k}(i)}\{\alpha_j^{-1}(\varpi_j)\}. \tag{6.2.22}$$

By setting $\vartheta_i = \max_{j\in\mathcal{N}_{I_k}(i)}\{\alpha_j^{-1}(\varpi_j) + \phi_{ij}\}$ and combining (6.2.22) with (6.2.21), one has,

$$\begin{aligned}\rho_{inti}(\vartheta_i) &= \rho_{inti}(\max_{j\in\mathcal{N}_{I_k}(i)}\{\alpha_j^{-1}(\varpi_j) + \phi_{ij}\})\\ &\leq \rho_{inti}(\max_{j\in\mathcal{N}_{I_k}(i)}\{\alpha_j^{-1}(\varpi_j)\} + \max_{j\in\mathcal{N}_{I_k}(i)}\{\phi_{ij}\}) < \kappa_i(\varpi_i),\end{aligned}$$

which again implies that one can always find suitable local parameters η_i, μ_i to satisfy (6.2.12) (resp. (6.2.15)). Additionally, the selection of $\vartheta_i = \max_{j\in\mathcal{N}_{I_k}(i)}\{\alpha_j^{-1}(\varpi_j)+\phi_{ij}\}$ as in lines 5 and 8, together with the design procedure for ϖ_i and ϕ_{ij} ensure that (6.2.7) is satisfied as well, which concludes the proof. □

Notice that the design procedure in Algorithm 1 follows the hierarchy of the acyclic directed graph which is composed of SCCs as vertices. Since the interconnected system considered in this section is composed of finite number of SCCs, Algorithm 1 terminates in finite iterations.

The compositionality scheme proposed here is schematically illustrated in Figure 6.3.

Remark 6.2.19. *As can be observed from the theorem, the compositional framework is based on a small-gain type condition. Small-gain theorems have a long-known history in control design dating back to the 1960's [219]. They have been extensively leveraged to establish stability properties of interconnected systems [78, 38]. In recent years, small-gain type conditions have been leveraged in [146, 126, 188, 184] to facilitate the compositional construction of finite abstractions. The results in [146, 126, 188] rely on classic sum-type small-gain conditions which require almost linear growth on gains of subsystems. In contrast, our compositionality result here are based on max-type small-gain conditions formulated in a general nonlinear form, which can potentially lead to much smaller approximation errors of finite abstractions; see [184, Remark 3.6] for some discussions on this point. It should be noted that if the small-gain type condition* (6.2.19) *is satisfied by every SCC in the network, then this condition holds for the*

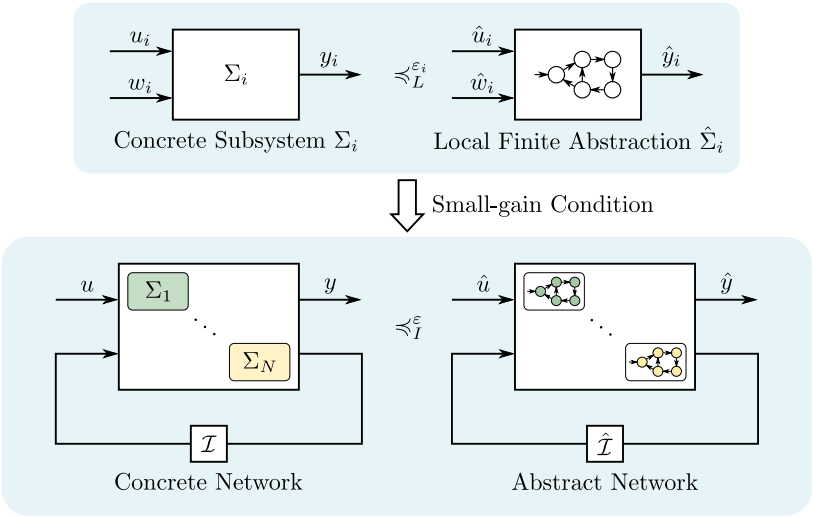


Figure 6.3: Compositional framework for the construction of opacity-preserving finite abstractions for interconnected systems.

overall network as well. However, by involving the notion of SCCs in the parameter design procedure, we are allowed to check the small-gain condition and design local parameters inside each SCC only, instead of the entire network. Moreover, by exploiting the interconnection topology, the proposed result presents a top-down *compositional design framework. That is, as long as Assumption 6.2.17 holds, given any desired precision $\varpi \in \mathbb{R}_{>0}$, Algorithm 1 always provides us with suitable local quantization parameters to achieve the overall abstraction accuracy. Note that such a systematic compositional scheme cannot be achieved by the results in [184].*

6.2.4 Case Study

6.2.4.1 Compositional Construction of Opacity-Preserving Finite Abstractions

Here, we provide an illustrative example to explain the design procedure of local quantization parameters using Algorithm 1. The system model is adapted from [146].

Consider the interconnected discrete-time system Σ consisting of $n = 6$ subsystems:

$$\Sigma : \begin{cases} \mathbf{x}_1(k+1) &= k_{11}\frac{\mathbf{x}_1(k)}{1+\mathbf{x}_1^2(k)} + \nu_1(k), \\ \mathbf{x}_2(k+1) &= k_{21}\tanh(\mathbf{x}_2(k)) + k_{22}(\operatorname{sech}(\mathbf{x}_3(k)) - 1 + \mathbf{x}_1(k)), \\ \mathbf{x}_3(k+1) &= k_{31}\mathbf{x}_3(k) + k_{32}(\sin\mathbf{x}_2(k) + \mathbf{x}_5(k)) + \nu_3(k), \\ \mathbf{x}_4(k+1) &= k_{41}(\cos(\mathbf{x}_4(k)) - 1) + k_{42}(\tanh(\mathbf{x}_5(k))), \\ \mathbf{x}_5(k+1) &= k_{51}\sin(\mathbf{x}_5(k)) + k_{52}(\operatorname{sech}(\mathbf{x}_4(k)) - 1) + \nu_5(k), \\ \mathbf{x}_6(k+1) &= k_{61}\frac{\mathbf{x}_6(k)}{1+|\mathbf{x}_6(k)|} + k_{62}\mathbf{x}_5(k), \\ \mathbf{y}(k) &= \mathbf{x}(k), \end{cases} \tag{6.2.23}$$

where $k \in \mathbb{N}$, $\mathbf{x}(k) = [\mathbf{x}_1(k); \ldots; \mathbf{x}_n(k)]$, $\mathbf{y}(k) = [\mathbf{x}_1(k); \ldots; \mathbf{x}_n(k)]$. The outputs of the subsystems are: $\mathbf{y}_i(k) = c_i\mathbf{x}_i(k)$, where $c_i = [c_{i1}; \ldots; c_{in}]$ with $c_1 = [1; 1; 0; 0; 0; 0]$,

$c_2 = [0;1;1;0;0;0]$, $c_3 = [0;1;1;0;0;0]$, $c_4 = [0;0;0;1;1;0]$, $c_5 = [0;0;1;1;1;1]$, $c_6 = [0;0;0;0;0;1]$, internal inputs subject to the constraints $w_i = [y_{1i}; \ldots; y_{(i-1)i}; y_{(i+1)i}; \ldots; y_{ni}]$, $\forall i \in [1;6]$, $\kappa_{i,1} = 0.4$, $\forall i \in [1;6]$, $\kappa_{i,2} = 0.2$, $\forall i \in [2;5]$, $X_i = [-1,1]$ and $U_i = [-1,1]$, $\forall i \in [1;6]$. One can readily verify that the system Σ in (6.2.23) can be seen as an interconnection of 6 scalar subsystems Σ_i, $i \in [1;6]$, as in Definition 6.2.2. The directed graph $G = (I, E)$ is specified by $I = [1;6]$, $E = \{(2,1),(3,2),(2,3),(5,4),(3,5),(4,5),(6,5)\}$. Strongly connected components of G are $\bar{G}_1$ with $I_1 = \{1\}$, $\bar{G}_2$ with $I_2 = \{4,5\}$, $\bar{G}_3$ with $I_3 = \{2,3\}$ and $\bar{G}_4$ with $I_4 = \{6\}$. Now we apply our main results in the previous sections to compositionally construct a finite abstraction of Σ with accuracy $\varepsilon = 0.01$ as defined in (6.2.6), which preserves approximate initial-state opacity.

First, let us choose functions $V_i = |x_i - x_i'|$, $\forall i \in [1;6]$. It can be readily seen that V_i are δ-ISS Lyapunov functions for subsystems Σ_i satisfying (6.2.9) and (6.2.10) in Definition 6.2.13, with $\kappa_i(s) = (1 - |\kappa_{i,1}|)s$, $\underline{\alpha}_i(s) = \overline{\alpha}_i(s) = \hat{\gamma}_i(s) = s$, $\rho_{int1}(s) = 0$, $\rho_{int2}(s) = 2|\kappa_{2,2}|s$, $\rho_{int3}(s) = 2|\kappa_{3,2}|s$, $\rho_{int4}(s) = |\kappa_{4,2}|s$, $\rho_{int5}(s) = |\kappa_{5,2}|s$, $\rho_{int6}(s) = |\kappa_{6,2}|s$, $\rho_{ext2}(s) = \rho_{ext4}(s) = \rho_{ext6}(s) = 0$, $\rho_{ext1}(s) = \rho_{ext3}(s) = \rho_{ext5}(s) = s$. The Lipschitz assumption holds with $\ell_i(s) = s$. Since we have $\gamma_{ij}(s) < \mathrm{id}$ as defined in (6.2.18), $\forall i, j \in I$, the small-gain condition (6.2.19) is readily satisfied for every SCC. Functions $\sigma_i = \mathrm{id}$, $\forall i \in I$, readily satisfy (6.2.20).

Now we apply Algorithm 1 to design the local parameters. The desired precision is $\varpi = 0.01$ by (6.2.6). We design for all of the subsystems, $\phi_i = 0$, $\forall i \in [1;6]$. We start with $G^* = G$ and get the associated $\mathrm{BSCC}(G^*) = \{\bar{G}_3, \bar{G}_4\}$ for line 3. First, let us consider the SCC $\bar{G}_3$. We choose $r = 0.01$ to satisfy the conditions in lines $8 - -9$ with $\varpi_2 = \vartheta_2 = \varpi_3 = \vartheta_3 = \varpi = 0.01$. For $\bar{G}_4$, since it contains only 1 subsystem Σ_6, we get in line 6, $\varpi_6 = \varpi = 0.01$ and choose $\vartheta_6 = 0.01$. Now G^* is updated in line 12 to $\{\bar{G}_1, \bar{G}_2\}$. The bottom SCCs of the updated G^* is $\{\bar{G}_1, \bar{G}_2\}$. Since the current graph G^* is not the entire network anymore, we go to lines $8 - -9$. We proceed with $\bar{G}_1$ firstly. Since $\bar{G}_1$ consists of only 1 subsystem, we go to line 9 and set $\varpi_1 = \vartheta_2 = 0.01$ and $\vartheta_1 = 0.01$ such that the inequalities hold. Now consider $\bar{G}_2$. In line 8, we choose $r = \min\{\vartheta_3, \vartheta_6\} = 0.01$, and then set $\varpi_4 = \varpi_5 = r$, $\vartheta_4 = \varpi_5$ and $\vartheta_5 = \varpi_4$ in line 8. Next, the set G^* becomes empty and the algorithm ends. Till now, we obtain local parameters (ϖ_i, ϑ_i) for each subsystem. Now we have the freedom to design the local quantization parameters η_i, μ_i using (ϖ_i, ϑ_i) while satisfying inequality (6.2.12). We show here a choice of suitable tuples of local parameters $q_i = (\eta_i, \theta_i, \mu_i, \phi_i)$ as: $q_1 = (0.006, 0, 0, 0)$, $q_2 = (0.002, 0, 0, 0)$, $q_3 = (0.002, 0, 0, 0)$, $q_4 = (0.004, 0, 0, 0)$, $q_5 = (0.004, 0, 0, 0)$, $q_6 = (0.004, 0, 0, 0)$. Now, one can construct local abstractions for subsystems as in Subsection 6.2.3.1. Using the result in Theorem 6.2.14, one can verify that $V_i = |x_i - x_i'|$ is a ϖ_i-InitSOPSF from each Σ_i to its abstraction $\hat{\Sigma}_i$. By the results in Theorem 6.2.12, one can verify that $\tilde{V}(x, \hat{x}) = \max_i\{|x_i - \hat{x}_i|\}$ is a ϖ-InitSOPSF from Σ to $\hat{\Sigma} = \mathcal{I}(\hat{\Sigma}_1, \ldots, \hat{\Sigma}_n)$.

6.2.4.2 Verification of Initial-State Opacity for An Interconnected System

Consider a concrete interconnected discrete-time linear system Σ, consisting of $n \in \mathbb{N}_{\geq 1}$ subsystems Σ_i, each described by:

$$\Sigma_i : \begin{cases} \mathbf{x}_i(k+1) &= a_i \mathbf{x}_i(k) + \nu_i(k) + d_i w_i(k), \\ \mathbf{y}_i(k) &= c_i \mathbf{x}_i(k), \end{cases}$$

where $a_i = 0.1$, $d_i = 0.05$, $c_i = [c_{i1}; \ldots; c_{in}]$ with $c_{i(i+1)} = 1$, $c_{ij} = 0$, $\forall i \in [1; n-1], \forall j \neq i+1$, $c_{nn} = 1$, $c_{nj} = 0$, $\forall j \in [1; n-1]$, $\nu_i(k) = 0.145$, $w_1(k) = 0$, and $w_i(k) = \mathbf{y}_{(i-1)i}(k)$, $\forall i \in [2; n]$. For each subsystem, the state set is $X_i = X_{0_i} =]0\ 0.6[$, the input set is $U_i = \{0.145\}$, the secret set is $X_{S_1} =]0\ 0.2]$, $X_{S_2} = [0.4\ 0.6[$, $X_{S_i} =]0\ 0.6[$, $\forall i \in [3; n]$, the output set is $Y_i = \prod_{j=1}^{n} Y_{ij}$ where $Y_{i(i+1)} =]0\ 0.6[$, $Y_{ij} = 0$, $\forall i \in [1; n-1]$, $\forall j \neq i+1$, $Y_{nn} =]0\ 0.6[$, $Y_{nj} = 0$, $\forall j \in [1; n-1]$, and the internal input set is $W_i = \prod_{j=1, j\neq i}^{n} Y_{ji}$. Intuitively, the output of the overall system is the external output of the last subsystem Σ_n. The main goal of this example is to verify approximate initial-state opacity of the concrete network using its finite abstraction. Now, let us construct compositionally a finite abstraction of Σ that preserves initial-state opacity, with desired accuracy $\varepsilon = 0.25$ in Proposition 6.2.8. We apply our main results of previous sections to achieve this goal. Consider functions $V_i = |x_i - x_i'|$, $\forall i \in [1; n]$. It can be is readily verified that V_i are δ-ISS Lyapunov functions for subsystems Σ_i satisfying (6.2.9) and (6.2.10) in Definition 6.2.13, with $\kappa_i(s) = (1 - a_i)s = 0.9s$, $\rho_{exti}(s) = \hat{\gamma}_i(s) = \underline{\alpha}_i(s) = \overline{\alpha}_i(s) = s$, and $\rho_{inti}(s) = 0.05s$. It can be seen that the system is made up of n identical subsystems in a cascade interconnection, thus, the resulting directed graph $G = (I, E)$ is specified by $I = [1; n]$, $E = \{(1, 2), (2, 3), (3, 4), \ldots, (n-1, n)\}$. Note that each subsystem is a strongly connected component of G and the small-gain condition (6.2.19) is satisfied readily. Then, by applying Algorithm 1 and choosing functions $\sigma_i = \mathrm{id}$, $\forall i \in [1; n]$, we obtain proper pairs of local parameters $(\varpi_i, \vartheta_i) = (0.25, 0.25)$ for all of the subsystems. Then, a suitable tuple $q_i = (\eta_i, \mu_i, \theta_i, \phi_i) = (0.2, 0, 0, 0)$ of quantization parameters is chosen such that inequality (6.2.12) for each subsystem Σ_i is satisfied. Next, we construct local abstractions $\hat{\Sigma}_i = (\hat{X}_i, \hat{X}_{0_i}, \hat{X}_{S_i}, \hat{U}_i, \hat{W}_i, \hat{f}_i, \hat{Y}_i, \hat{h}_i)$ for subsystems as in Subsection 6.2.3.1, where $\hat{X}_i = \hat{X}_{0_i} = \{0.2, 0.4\}$, $\hat{X}_{S_1} = \{0.2\}$, $\hat{X}_{S_2} = \{0.4\}$, $\hat{X}_{S_i} = \{0.2, 0.4\}$,$\forall i \in [3; n]$, $\hat{Y}_i = \prod_{j=1}^{i}\{0\} \times \{0.2, 0.4\} \times \prod_{j=i+2}^{n}\{0\}$, $\forall i \in [1; n-1]$, $\hat{Y}_n = \prod_{j=1}^{n-1}\{0\} \times \{0.2, 0.4\}$, $\hat{W}_i = \{0.2, 0.4\}$, $\forall i \in [1; n]$. Using the result in Theorem 6.2.14, one can verify that $V_i = |x_i - x_i'|$ is a local ϖ_i-InitSOPSF from each Σ_i to its abstraction $\hat{\Sigma}_i$. Furthermore, by the compositionality result in Theorem 6.2.12, we obtain that $\tilde{V} = \max_i\{V_i(x_i, \hat{x}_i)\} = \max_i\{|x_i - x_i'|\}$ is an ϖ-InitSOPSF from $\Sigma = \mathcal{I}(\Sigma_1, \ldots, \Sigma_n)$ to $\hat{\Sigma} = \hat{\mathcal{I}}(\hat{\Sigma}_1, \ldots, \hat{\Sigma}_n)$ satisfying the conditions in Definition 6.2.4 with $\varpi = \max_i \varpi_i = 0.25$.

Now, let us verify approximate initial-state opacity for Σ using the interconnected abstraction $\hat{\Sigma}$. An example of a network consisting of 3 subsystems is shown in Figure 6.4. The three smaller automata in the left represent the symbolic subsystems and the one in the right represents the interconnected abstraction for the whole network. For simplicity of demonstration, we use symbols to represent the state and output vectors, where the states and outputs of local transition systems are denoted by $a = [0.2]$,

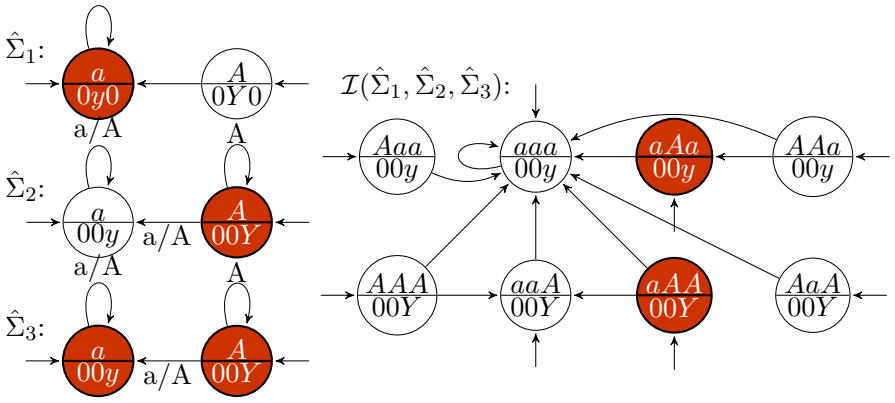


Figure 6.4: Compositional abstraction of an interconnected discrete-time linear system consisting of 3 subsystems. Each circle is labeled by the state (top half) and the corresponding output (bottom half). Initial states are distinguished by being the target of a sourceless arrow. Secret states are marked in red. The symbols on the edges show the internal inputs coming from other subsystems.

$A = [0.4]$, $y = 0.2$ and $Y = 0.4$, respectively. The symbols such as $aaa = [0.2; 0.2; 0.2]$ and $00y = [0; 0; 0.2]$ represent the concatenated state and output vectors for the interconnected abstraction, respectively. As seen in Figure 6.4, for any run starting from any secret state, i.e., aAa and aAA, there exists a run from a non-secret state, i.e., Aaa and AAA, such that the output trajectories are exactly the same. Due to lack of space, we do not plot the automata for the case of $n = 4$, but we verified that the network is still 0-approximate initial-state opaque. We expect that the network holds this property regardless of the number of subsystems due to the homogeneity of subsystems and the structure of the network topology. Thus, one can conclude that $\hat{\mathcal{I}}(\hat{\Sigma}_1, \ldots, \hat{\Sigma}_n)$ is 0-approximate initial-state opaque. Therefore, by Proposition 6.2.8, we obtain that the concrete network $\mathcal{I}(\Sigma_1, \ldots, \Sigma_n)$ is 0.5-approximate initial-state opaque.

6.3 An Abstraction-based Approach for Interconnected Switched Systems

The results in the previous section present a compositional framework for the construction of opacity-preserving finite abstractions for interconnected *control* systems without any discrete dynamic. In this section, we enlarge the class of systems for the first time to hybrid ones with switching signals.

First, we introduce the definitions of discrete-time interconnected switched systems and subsystems. Then, new notions of approximate opacity-preserving simulation functions are proposed. Next, we provide a compositional framework for the construction of opacity-preserving simulation functions for a network of discrete-time switched systems. Finally, we present how to construct local finite abstractions for a class of incremen-

tally input-to-state stable subsystems, and then propose a small-gain type condition required for the main compositionality result.

6.3.1 Interconnected Switched Systems

6.3.1.1 Discrete-Time Switched Subsystems

We consider discrete-time switched subsystems of the following form.

Definition 6.3.1. *A discrete-time switched subsystem Σ is defined by the tuple $\Sigma = (\mathbb{X}, \mathbb{X}_0, \mathbb{X}_S, P, W, F, Y, h)$, where*

- $\mathbb{X} \subseteq \mathbb{R}^n$ *is the state set;*
- $\mathbb{X}_0 \subseteq X$ *is the initial state set;*
- $\mathbb{X}_S \subseteq X$ *is the secret state set;*
- $P = \{1, \ldots, m\}$ *is the finite set of modes;*
- $W \subseteq \mathbb{R}^m$ *is the internal input set;*
- $F = \{f_1, \ldots, f_m\}$ *is a collection of set-valued maps $f_p : X \times W \rightrightarrows X$ for all $p \in P$;*
- $Y \subseteq \mathbb{R}^q$ *is the output set;*
- $h : X \to Y$ *is the output map.*

The discrete-time switched subsystem Σ is described by difference inclusions of the form

$$\Sigma : \begin{cases} \mathbf{x}(k+1) & \in f_{\mathsf{p}(k)}(\mathbf{x}(k), \omega(k)), \\ \mathbf{y}(k) & = h(\mathbf{x}(k)), \end{cases} \tag{6.3.1}$$

where $\mathbf{x} : \mathbb{N} \to \mathbb{X}$, $\mathbf{y} : \mathbb{N} \to Y$, $\mathsf{p} : \mathbb{N} \to P$, and $\omega : \mathbb{N} \to W$ are the state, output, switching, and internal input signal, respectively.

Let $\varphi_k, k \in \mathbb{N}_{\geq 1}$, denote the time when the k-th switching instant occurs. We assume that signal p satisfies a dwell-time condition [106] (i.e. there exists $k_d \in \mathbb{N}_{\geq 1}$, called the dwell-time, such that for all consecutive switching time instants φ_k, φ_{k+1}, $\varphi_{k+1} - \varphi_k \geq k_d$). If for all $x \in \mathbb{X}, p \in P, w \in W$, $\mathrm{card}(f_p(x, w)) \leq 1$, we say the system Σ is deterministic, and non-deterministic otherwise. System Σ is called finite if X, W are finite sets and infinite otherwise. We assume that for every initial condition and any sequence of switching signals, the corresponding state signal is defined for all $k \geq 0$.

Similar as in Definition 2.3.4, we employ the notion of transition systems to provide an alternative description of switched systems that can be later directly related to their finite abstractions in a common framework.

Definition 6.3.2. *Given a discrete-time switched subsystem $\Sigma = (\mathbb{X}, \mathbb{X}_0, \mathbb{X}_S, P, W, F, Y, h)$, we define the associated transition system $T(\Sigma) = (X, X_0, X_S, U, W, \mathcal{F}, Y, \mathcal{H})$, where:*

- $X = \mathbb{X} \times P \times \{0, \dots, k_d - 1\}$ *is the state set;*
- $X_0 = \mathbb{X}_0 \times P \times \{0\}$ *is the initial state set;*
- $X_S = \mathbb{X}_S \times P \times \{0, \dots, k_d - 1\}$ *is the secret state set;*
- $U = P$ *is the external input set;*
- $W = W$ *is the internal input set;*
- $\mathcal{F}$ *is the transition function given by* $(x^+, p^+, l^+) \in \mathcal{F}((x, p, l), u, w)$ *if and only if* $x^+ \in f_p(x, w), u = p$ *and one of the following scenarios hold:*
  - $l < k_d - 1$, $p^+ = p$ *and* $l^+ = l + 1$*: switching is not allowed because the time elapsed since the latest switch is strictly smaller than the dwell time;*
  - $l = k_d - 1$, $p^+ = p$ *and* $l^+ = k_d - 1$*: switching is allowed but no switch occurs;*
  - $l = k_d - 1$, $p^+ \neq p$ *and* $l^+ = 0$*: switching is allowed and a switch occurs;*
- $Y = Y$ *is the output set;*
- $\mathcal{H} : X \to Y$ *is the output map defined as* $\mathcal{H}(x, p, l) = h(x)$.

Note that in the above definition, two additional variables p and l are added to the state tuple of the system Σ. The variable l serves as a counter to record the sojourn time of the switching signal, which allows or prevents the system from switching depending on whether the dwell-time condition is satisfied; the variable p acts as a memory to record the current mode of the system.

The following proposition is borrowed from [184] showing that the output runs of a discrete-time switched subsystem Σ and its associated transition system $T(\Sigma)$ are equivalent so that one can use Σ and $T(\Sigma)$ interchangeably.

Proposition 6.3.3. *Consider a transition system $T(\Sigma)$ in Definition 6.3.2 associated to Σ in Definition 6.3.1. Any output trajectory of Σ can be uniquely matched to an output trajectory of $T(\Sigma)$ and vice versa.*

Next, let us introduce a formal definition of networks of dt-SS (or equivalently, networks of transition systems).

6.3.1.2 Discrete-Time Interconnected Switched Systems

Consider $N \in \mathbb{N}_{\geq 1}$ discrete-time switched subsystems $\Sigma_i = (\mathbb{X}_i, \mathbb{X}_{0_i}, \mathbb{X}_{S_i}, P_i, W_i, F_i, Y_i, h_i)$, $i \in [1; N]$, with partitioned internal inputs and outputs as

$$w_i = [w_{i1}; \dots; w_{i(i-1)}; w_{i(i+1)}; \dots; w_{iN}], \quad W_i = \prod_{j=1, j\neq i}^{N} W_{ij}, \tag{6.3.2}$$

$$h_i(x_i) = [h_{i1}(x_i); \ldots; h_{iN}(x_i)], \quad Y_i = \prod_{j=1}^{N} Y_{ij}, \tag{6.3.3}$$

with $w_{ij} \in W_{ij}$, and $y_{ij} = h_{ij}(x_i) \in Y_{ij}$. The outputs y_{ii} are considered as external ones, whereas y_{ij} with $i \neq j$ are interpreted as internal ones which are used to construct interconnections between systems. In particular, we assume that $w_{ij} = y_{ji}$, if there is connection from system Σ_j to Σ_i, otherwise, we set $h_{ji} \equiv 0$. In the sequel, we denote by $\mathcal{N}_i = \{j \in [1; N], j \neq i | h_{ji} \neq 0\}$ the collection of neighboring subsystems $\Sigma_j, j \in \mathcal{N}_i$, that provide internal inputs to subsystem Σ_i.

Now, we introduce the notion of networks (in both concrete and abstract domains) based on the notion of interconnected systems in [188]. For a concrete network constructed as the interconnection of $N \in \mathbb{N}_{\geq 1}$ concrete subsystems, Definition 6.3.1 reduces to the tuple $\Sigma = (\mathbb{X}, \mathbb{X}_0, \mathbb{X}_S, P, F, Y, h)$ without internal inputs and outputs as in the following definition. Note that an interconnected switched system without internal inputs and outputs reduces to a discrete-time switched system as in Definition 2.3.3.

Definition 6.3.4. *Consider $N \in \mathbb{N}_{\geq 1}$ switched subsystems $\Sigma_i = (\mathbb{X}_i, \mathbb{X}_{0_i}, \mathbb{X}_{S_i}, P_i, W_i, F_i, Y_i, h_i)$, $i \in [1; N]$ with the input-output structure given by* (6.3.2) *and* (6.3.3)*. The network, representing the interconnection of $N \in \mathbb{N}_{\geq 1}$ switched subsystems Σ_i, is a tuple $\Sigma = (\mathbb{X}, \mathbb{X}_0, \mathbb{X}_S, P, F, Y, h)$, denoted by $\mathcal{I}(\Sigma_1, \ldots, \Sigma_N)$, where $\mathbb{X} = \prod_{i=1}^{N} \mathbb{X}_i$, $\mathbb{X}_0 = \prod_{i=1}^{N} \mathbb{X}_{0_i}$, $\mathbb{X}_S = \prod_{i=1}^{N} \mathbb{X}_{S_i}$, $P = \prod_{i=1}^{N} P_i$, $F = \prod_{i=1}^{N} F_i$, $Y = \prod_{i=1}^{N} Y_{ii}$, $h(x) := [h_{11}(x_1); \ldots; h_{NN}(x_N)]$ with $x = [x_1; \ldots; x_N]$, subject to the constraint:*

$$y_{ji} = w_{ij}, Y_{ji} \subseteq W_{ij}, \forall i \in [1; N], j \in \mathcal{N}_i. \tag{6.3.4}$$

Similarly, given transition systems $T(\Sigma_i)$, one can also define a network of transition systems $\mathcal{I}(T(\Sigma_1), \ldots, T(\Sigma_N))$. For the rest of the section, we mainly deal with the transition systems as they allow us to model switched subsystems Σ and their finite abstractions in a common framework.

For an interconnection of $N \in \mathbb{N}_{\geq 1}$ finite discrete-time switched subsystems $\hat{\Sigma}_i$, with input-output structure configuration as in (6.3.2) and (6.3.3), we introduce the following definition of networks of finite discrete-time switched subsystems.

Definition 6.3.5. *Consider $N \in \mathbb{N}_{\geq 1}$ finite switched subsystems $\hat{\Sigma}_i = (\hat{\mathbb{X}}_i, \hat{\mathbb{X}}_{0_i}, \hat{\mathbb{X}}_{S_i}, \hat{P}_i, \hat{W}_i, \hat{F}_i, \hat{Y}_i, \hat{h}_i)$, $i \in [1; N]$ with the input-output structure given by* (6.3.2) *and* (6.3.3)*. The network, representing the interconnection of $N \in \mathbb{N}_{\geq 1}$ finite switched subsystems $\hat{\Sigma}_i$, is a tuple $\hat{\Sigma} = (\hat{\mathbb{X}}, \hat{\mathbb{X}}_0, \hat{\mathbb{X}}_S, \hat{P}, \hat{F}, \hat{Y}, \hat{h})$, denoted by $\hat{\mathcal{I}}(\hat{\Sigma}_1, \ldots, \hat{\Sigma}_N)$, where $\hat{\mathbb{X}} = \prod_{i=1}^{N} \hat{\mathbb{X}}_i$, $\hat{\mathbb{X}}_0 = \prod_{i=1}^{N} \hat{\mathbb{X}}_{0_i}$, $\hat{\mathbb{X}}_S = \prod_{i=1}^{N} \hat{\mathbb{X}}_{S_i}$, $\hat{P} = \prod_{i=1}^{N} \hat{P}_i$, $\hat{F} = \prod_{i=1}^{N} \hat{F}_i$, $\hat{Y} = \prod_{i=1}^{N} \hat{Y}_{ii}$, $\hat{h}(x) := \left[\hat{h}_{11}(\hat{x}_1); \ldots; \hat{h}_{NN}(\hat{x}_N)\right]$ with $\hat{x} = [\hat{x}_1; \ldots; \hat{x}_N]$, subject to the constraint:*

$$\|\hat{y}_{ji} - \hat{w}_{ij}\| \leq \phi_{ij}, [\hat{Y}_{ji}]_{\phi_{ij}} \subseteq \hat{W}_{ij}, \forall i \in [1; N], j \in \mathcal{N}_i, \tag{6.3.5}$$

where ϕ_{ij} is an internal input quantization parameter designed for constructing local finite abstractions (cf. Definition 6.3.20).

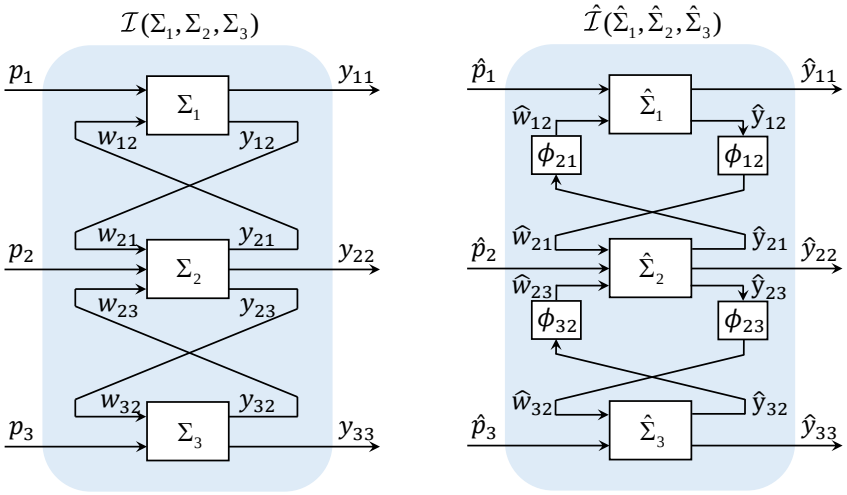


Figure 6.5: [Left]: Concrete network composed of three switched subsystems Σ_1, Σ_2, and Σ_3 with $h_{13} = h_{31} = 0$, where $y_{ji} = w_{ij}$, $\forall i, j \in [1;3]$; [Right]: Abstract network composed of three finite subsystems $\hat{\Sigma}_1$, $\hat{\Sigma}_2$, and $\hat{\Sigma}_3$ with $\hat{h}_{13} = \hat{h}_{31} = 0$, and the internal inputs $\hat{w}_{ij}$ for system $\hat{\Sigma}_i$ are taken from the discretized internal outputs of system $\hat{\Sigma}_j$ under the constraint $\|\hat{y}_{ji} - \hat{w}_{ij}\| \leq \phi_{ij}$, $\forall i, j \in [1;3]$, where ϕ_{ij} are internal input quantization parameters.

Similarly, given finite transition systems $T(\hat{\Sigma}_i)$, one can also define a network of transition systems as $\hat{\mathcal{I}}(T(\hat{\Sigma}_1), \ldots, T(\hat{\Sigma}_N))$.

An example of a concrete network and an abstract network is illustrated in Figure 6.5, where each consists of three switched subsystems.

Remark 6.3.6. *Note that in the above definitions, the interconnection constraint in* (6.3.4) *for the concrete network is different from that for the abstract network in* (6.3.5)*. For networks of finite abstractions, due to possibly different granularities of finite internal input sets $\hat{W}_{ij}$ and output sets $\hat{Y}_{ij}$, we introduce parameters ϕ_{ij} in* (6.3.5) *for having a well-posed interconnection. The values of ϕ_{ij} will be designed later in Definition 6.3.20 when constructing local finite abstractions of the subsystems.*

We introduce some notations that will be used to characterize opacity property. Consider network $T(\Sigma)$. We use z^k to denote a state of $T(\Sigma)$ reached at time $k \in \mathbb{N}$ from initial state z^0 under an input sequence $\bar{u}$ with length k, and denote by $\{z^0, z^1, \ldots, z^n\}$ a finite state run of $T(\Sigma)$ with length $n \in \mathbb{N}$.

For the sake of clarity, let us recall the notion of approximate initial-state (resp. current-state) opacity (cf. Definition 3.3.1) which is rewritten here to fit in the context of discrete-time interconnected switched systems.

Definition 6.3.7. *Consider network $T(\Sigma) = (X, X_0, X_S, U, \mathcal{F}, Y, \mathcal{H})$ and a constant $\delta \geq 0$. Network $T(\Sigma)$ is said to be*

- δ-approximate initial-state opaque *if for any* $z^0 \in X_0 \cap X_S$ *and any finite state run* $\{z^0, z^1, \ldots, z^n\}$, *there exist* $\bar{z}^0 \in X_0 \setminus X_S$ *and a finite state run* $\{\bar{z}^0, \bar{z}^1, \ldots, \bar{z}^n\}$ *such that* $\max_{k\in[0;n]} \|\mathcal{H}(z^k) - \mathcal{H}(\bar{z}^k)\| \leq \delta$.
- δ-approximate current-state opaque *if for any* $z^0 \in X_0$ *and finite state run* $\{z^0, z^1, \ldots, z^n\}$ *such that* $z^n \in X_S$, *there exists* $\bar{z}^0 \in X_0$ *and a finite state run* $\{\bar{z}^0, \bar{z}^1, \ldots, \bar{z}^n\}$ *such that* $\bar{z}^n \in X \setminus X_S$ *and* $\max_{k\in[0;n]} \|\mathcal{H}(z^k) - \mathcal{H}(\bar{z}^k)\| \leq \delta$.

In the next corollary, we show that if a system equipped with secret set X_S is δ-approximate opaque, then the system is also δ-approximate opaque with a smaller secret set contained in X_S.

Corollary 6.3.8. *Consider networks* $T(\Sigma_1) = (X, X_0, X_S, U, \mathcal{F}, Y, \mathcal{H})$ *and* $T(\Sigma_2) = (X, X_0, X'_S, U, \mathcal{F}, Y, \mathcal{H})$ *with* $X'_S \subseteq X_S$. *If* $T(\Sigma_1)$ *is* δ*-approximate initial-state (resp. current-state) opaque, then* $T(\Sigma_2)$ *is also* δ*-approximate initial-state (resp. current-state) opaque.*

Proof. We start by showing the preservation of approximate initial-state opacity across systems $T(\Sigma_1)$ and $T(\Sigma_2)$. Consider any $z^0 \in X_0 \cap X'_S$ and any finite state run $\{z^0, z^1, \ldots, z^n\}$ in $T(\Sigma_2)$. Given that $X'_S \subseteq X_S$, we get $z^0 \in X_0 \cap X_S$. Since $T(\Sigma_1)$ is δ-approximate initial-state opaque, from Definition 6.3.7, there exist $\bar{z}^0 \in X_0 \setminus X_S$ and a finite state run $\{\bar{z}^0, \bar{z}^1, \ldots, \bar{z}^n\}$ such that $\max_{k\in[0;n]} \|\mathcal{H}(z^k) - \mathcal{H}(\bar{z}^k)\| \leq \delta$. Moreover, given that $\{X_0 \setminus X_S\} \subseteq \{X_0 \setminus X'_S\}$, we get $\bar{z}^0 \in X_0 \setminus X'_S$. Therefore, by Definition 6.3.7, $T(\Sigma_2)$ is also δ-approximate initial-state opaque. Similarly, we show the preservation of approximate current-state opacity across systems $T(\Sigma_1)$ and $T(\Sigma_2)$. Consider any $z^0 \in X_0$ and any finite state run $\{z^0, z^1, \ldots, z^n\}$ such that $z^n \in X'_S$ in $T(\Sigma_2)$. Since $T(\Sigma_1)$ is δ-approximate current-state opaque, from Definition 6.3.7, there exist $\bar{z}^0 \in X_0$ and a finite state run $\{\bar{z}^0, \bar{z}^1, \ldots, \bar{z}^n\}$ such that $\bar{z}^n \in X \setminus X_S$ and $\max_{k\in[0;n]} \|\mathcal{H}(z^k) - \mathcal{H}(\bar{z}^k)\| \leq \delta$. Moreover, given that $\{X \setminus X_S\} \subseteq \{X \setminus X'_S\}$, we get $\bar{z}^n \in X \setminus X'_S$. Therefore, by Definition 6.3.7, $T(\Sigma_2)$ is also δ-approximate current-state opaque.

□

Remark 6.3.9. *Note that it is assumed in Definitions 6.3.4 and 6.3.5 that the secret set of the network is the Cartesian product of the secret sets of the subsystems. However, if the secret set of the original network is in a more general from (e.g. polytopes), one can use minimum bounding box algorithms [15] to compute the smallest hyper-rectangle containing the secret set of the original network. If we consider this hyper-rectangle as the new secret set and follow the same procedure to verify opacity of the system, then by Corollary 6.3.8, the results (if successful) can be carried over to the original network.*

Remark 6.3.10. *The above-mentioned algorithms helps us to verify opacity for networks consisting of finite abstractions and then carry back the verification result to concrete ones, given a formal simulation relation between those networks. To this purpose, an opacity-preserving simulation relation will be introduced in the next section which formally relate a network of transition systems and its finite abstraction.*

6.3.2 Opacity-Preserving Simulation Functions

In this section, we introduce notions of approximate opacity-preserving simulation functions to quantitatively relate two networks of transition systems in terms of preserving approximate initial-state and current-state opacity. Such a function can be constructed compositionally as shown in Section 6.3.3.

6.3.2.1 Opacity-Preserving Simulation Functions

First, we introduce a notion of approximate initial-state opacity-preserving simulation functions in the following definition.

Definition 6.3.11. *Consider networks* $T(\Sigma) = (X, X_0, X_S, U, \mathcal{F}, Y, \mathcal{H})$ *and* $T(\hat{\Sigma}) = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{\mathcal{F}}, \hat{Y}, \hat{\mathcal{H}})$ *with* $\hat{Y} \subseteq Y$*. For* $\varepsilon \in \mathbb{R}_{\geq 0}$*, a function* $\mathcal{S} : X \times \hat{X} \to \mathbb{R}_{\geq 0}$ *is called an* ε*-approximate initial-state opacity-preserving simulation function (*ε*-InitSOPSF from* $T(\Sigma)$ *to* $T(\hat{\Sigma})$ *if there exists a function* $\alpha \in \mathcal{K}_\infty$ *such that*

1. *a)* $\forall z^0 \in X_0 \cap X_S$, $\exists \hat{z}^0 \in \hat{X}_0 \cap \hat{X}_S$, *s.t.* $\mathcal{S}(z^0, \hat{z}^0) \leq \varepsilon$;
   b) $\forall \hat{z}^0 \in \hat{X}_0 \setminus \hat{X}_S$, $\exists z^0 \in X_0 \setminus X_S$, *s.t.* $\mathcal{S}(z^0, \hat{z}^0) \leq \varepsilon$;
2. $\forall z \in X, \forall \hat{z} \in \hat{X}$, $\alpha(\|\mathcal{H}(z) - \hat{\mathcal{H}}(\hat{z})\|) \leq \mathcal{S}(z, \hat{z})$;
3. $\forall z \in X, \forall \hat{z} \in \hat{X}$ *s.t.* $\mathcal{S}(z, \hat{z}) \leq \varepsilon$, *one has:*
   a) $\forall u \in U$, $\forall z^+ \in \mathcal{F}(z, u)$, $\exists \hat{u} \in \hat{U}$, $\exists \hat{z}^+ \in \hat{\mathcal{F}}(\hat{z}, \hat{u})$, *s.t.* $\mathcal{S}(z^+, \hat{z}^+) \leq \varepsilon$;
   b) $\forall \hat{u} \in \hat{U}$, $\forall \hat{z}^+ \in \hat{\mathcal{F}}(\hat{z}, \hat{u})$, $\exists u \in U$, $\exists z^+ \in \mathcal{F}(z, u)$, *s.t.* $\mathcal{S}(z^+, \hat{z}^+) \leq \varepsilon$.

Similarly, we introduce a notion of approximate current-state opacity-preserving simulation functions defined as follows.

Definition 6.3.12. *Consider networks* $T(\Sigma) = (X, X_0, X_S, U, \mathcal{F}, Y, \mathcal{H})$ *and* $T(\hat{\Sigma}) = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{\mathcal{F}}, \hat{Y}, \hat{\mathcal{H}})$ *with* $\hat{Y} \subseteq Y$*. For* $\varepsilon \in \mathbb{R}_{\geq 0}$*, a function* $\mathcal{S} : X \times \hat{X} \to \mathbb{R}_{\geq 0}$ *is called an* ε*-approximate current-state opacity-preserving simulation function (*ε*-CurSOPSF) from* $T(\Sigma)$ *to* $T(\hat{\Sigma})$ *if there exists a function* $\alpha \in \mathcal{K}_\infty$ *such that*

1. $\forall z^0 \in X_0$, $\exists \hat{z}^0 \in \hat{X}_0$, *s.t.* $\mathcal{S}(z^0, \hat{z}^0) \leq \varepsilon$;
2. $\forall z \in X, \forall \hat{z} \in \hat{X}$, $\alpha(\|\mathcal{H}(z) - \hat{\mathcal{H}}(\hat{z})\|) \leq \mathcal{S}(z, \hat{z})$;
3. $\forall z \in X, \forall \hat{z} \in \hat{X}$ *s.t.* $\mathcal{S}(z, \hat{z}) \leq \varepsilon$, *one has:*
   a) $\forall u \in U$, $\forall z^+ \in \mathcal{F}(z, u)$, $\exists \hat{u} \in \hat{U}$, $\exists \hat{z}^+ \in \hat{\mathcal{F}}(\hat{z}, \hat{u})$, *s.t.* $\mathcal{S}(z^+, \hat{z}^+) \leq \varepsilon$;
   b) $\forall u \in U$, $\forall z^+ \in \mathcal{F}(z, u)$ *s.t.* $z^+ \in X_S$, $\exists \hat{u} \in \hat{U}$, $\exists \hat{z}^+ \in \hat{\mathcal{F}}(\hat{z}, \hat{u})$ *with* $\hat{z}^+ \in \hat{X}_S$, *s.t.* $\mathcal{S}(z^+, \hat{z}^+) \leq \varepsilon$;
   c) $\forall \hat{u} \in \hat{U}$, $\forall \hat{z}^+ \in \hat{\mathcal{F}}(\hat{z}, \hat{u})$ $\exists u \in U$, $\exists z^+ \in \mathcal{F}(z, u)$, *s.t.* $\mathcal{S}(z^+, \hat{z}^+) \leq \varepsilon$;
   d) $\forall \hat{u} \in \hat{U}$, $\forall \hat{z}^+ \in \hat{\mathcal{F}}(\hat{z}, \hat{u})$ *s.t.* $\hat{z}^+ \in \hat{X} \setminus \hat{X}_S$, $\exists u \in U$, $\exists z^+ \in \mathcal{F}(z, u)$ *with* $z^+ \in X \setminus X_S$, *s.t.* $\mathcal{S}(z^+, \hat{z}^+) \leq \varepsilon$.

We say that $T(\hat{\Sigma})$ is an abstraction of $T(\Sigma)$ if there exists an ε-InitSOPSF, or ε-CurSOPSF, from $T(\Sigma)$ to $T(\hat{\Sigma})$. In addition, if $T(\hat{\Sigma})$ is finite ($\hat{X}$ is a finite set), system $T(\hat{\Sigma})$ is called a finite abstraction (symbolic model) of the network $T(\Sigma)$, and is denoted by $T(\Sigma) \preceq^{\varepsilon} T(\hat{\Sigma})$.

Although Definitions 6.3.11 and 6.3.12 are general in the sense that networks $T(\Sigma)$ and $T(\hat{\Sigma})$ can be either infinite or finite, network $T(\hat{\Sigma})$ practically consists of $N \in \mathbb{N}_{\geq 1}$ finite abstractions. Hence, checking approximate initial-state, or current-state, opacity for the concrete network $T(\Sigma)$ can be done by resorting to that of its finite abstraction $T(\hat{\Sigma})$ and then carry the results back to the concrete network.

The next proposition shows that the existence of an ε-InitSOPSF (resp. ε-CurSOPSF) as we proposed in Definition 6.3.11 (resp. Definition 6.3.12) for networks of transition systems implies the existence of an approximate initial-state (resp. current-state) opacity-preserving simulation relations as in Definition 4.3.1 (resp. 4.3.6).

Proposition 6.3.13. *Consider networks $T(\Sigma) = (X, X_0, X_S, U, \mathcal{F}, Y, \mathcal{H})$ and $T(\hat{\Sigma}) = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{\mathcal{F}}, \hat{Y}, \hat{\mathcal{H}})$ where $\hat{Y} \subseteq Y$. Assume $\mathcal{S}$ is an ε-InitSOPSF (resp. CurSOPSF) from $T(\Sigma)$ to $T(\hat{\Sigma})$ as in Definition 6.3.11 (resp. Definition 6.3.12). Then, relation $R \subseteq X \times \hat{X}$ defined by*

$$R = \left\{(z, \hat{z}) \in X \times \hat{X} | \mathcal{S}(z, \hat{z}) \leq \varepsilon\right\}, \tag{6.3.6}$$

is an $\hat{\varepsilon}$-InitSOP (resp. $\hat{\varepsilon}$-CurSOP) simulation relation from $T(\Sigma)$ to $T(\hat{\Sigma})$ with

$$\hat{\varepsilon} = \alpha^{-1}(\varepsilon). \tag{6.3.7}$$

The proof of this proposition follows the same reasoning as that of Proposition 6.2.7.

Instead of directly working with the opacity-preserving simulation relations Definitions 4.3.1 and 4.3.6, in the sequel, we will mainly focus on the proposed notions of ε-InitSOPSFs and ε-CurSOPSFs as in Definitions 6.3.11 and 6.3.12 which allow us to establish our compositionality result in an easier way. We provide the following corollary which shows the usefulness of an approximate opacity-preserving simulation function in terms of preserving approximate opacity across related networks. The proof follows the same reasoning as that of Proposition 6.2.8 and is omitted here.

Corollary 6.3.14. *Consider networks $T(\Sigma) = (X, X_0, X_S, U, \mathcal{F}, Y, \mathcal{H})$ and $T(\hat{\Sigma}) = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{\mathcal{F}}, \hat{Y}, \hat{\mathcal{H}})$ where $\hat{Y} \subseteq Y$. Assume there exists an approximate opacity-preserving simulation function from $T(\Sigma)$ to $T(\hat{\Sigma})$ as in Definitions 6.3.11 and 6.3.12 associated with $\varepsilon \in \mathbb{R}_{\geq 0}$ and $\alpha \in \mathcal{K}_{\infty}$. Let $\hat{\varepsilon}, \delta \in \mathbb{R}_{\geq 0}$ where $\hat{\varepsilon} = \alpha^{-1}(\varepsilon)$ and $\hat{\varepsilon} \leq \frac{\delta}{2}$. Then the following implication holds*

$$\begin{aligned} &T(\hat{\Sigma}) \text{ is } (\delta - 2\hat{\varepsilon})\text{-approximate opaque} \\ &\Rightarrow T(\Sigma) \text{ is } \delta\text{-approximate opaque.} \end{aligned}$$

Note that the above implication across two related systems holds for both notions of approximate initial-state and current-state opacity in Definition 6.3.7. Corollary 6.3.14 provides us a sufficient condition for verifying approximate opacity of a complex network using abstraction-based techniques.

6.3.2.2 Compositional Construction of Opacity-Preserving Simulation Functions

As shown in the previous section, the proposed ε-InitSOPSF (resp. ε-CurSOPSF) can be used for checking approximate initial-state (resp. current-state) opacity of concrete networks by leveraging their finite abstractions. However, for a network consisting of a large number of switched subsystems, constructing the corresponding simulation function and the abstract network monolithically is not feasible in general due to curse of dimensionality. Hence, in this section, we introduce a compositional framework for the construction of opacity-preserving finite abstractions for networks of switched systems. In particular, we first relate local finite abstractions of the subsystems via local InitSOPSFs or CurSOPSFs. Then, one can obtain the abstract network by interconnecting the local finite abstractions of the subsystems. Additionally, the corresponding ε-InitSOPSF (resp. ε-CurSOPSF) to capture the closeness between the concrete and the abstract networks can be established by composing the local InitSOPSFs (resp. CurSOPSFs) as well.

Let us first introduce new notions of local InitSOPSFs and CurSOPSFs for switched subsystems with internal inputs in the following subsection.

Suppose that we are given N switched subsystems $\Sigma_i = (X_i, X_{0_i}, X_{S_i}, P_i, W_i, F_i, Y_i, h_i)$, $i \in [1; N]$, or equivalently, transition systems $T(\Sigma_i) = (X_i, X_{0_i}, X_{S_i}, U_i, W_i, \mathcal{F}_i, Y_i, \mathcal{H}_i)$. Moreover, we assume that each system $T(\Sigma_i)$ and its abstraction $T(\hat{\Sigma}_i)$ admit a local approximate opacity-preserving simulation function as defined next.

Definition 6.3.15. *Consider transition systems $T(\Sigma_i) = (X_i, X_{0_i}, X_{S_i}, U_i, W_i, \mathcal{F}_i, Y_i, \mathcal{H}_i)$ and $T(\hat{\Sigma}_i) = (\hat{X}_i, \hat{X}_{0_i}, \hat{X}_{S_i}, \hat{U}_i, \hat{W}_i, \hat{\mathcal{F}}_i, \hat{Y}_i, \hat{\mathcal{H}}_i)$, for all $i \in [1; N]$, where $\hat{W}_i \subseteq W_i$ and $\hat{Y}_i \subseteq Y_i$. For $\varepsilon_i \in \mathbb{R}_{\geq 0}$, a function $\mathcal{S}_i : X_i \times \hat{X}_i \to \mathbb{R}_{\geq 0}$ is called a local ε_i-InitSOPSF from $T(\Sigma_i)$ to $T(\hat{\Sigma}_i)$ if there exist a constant $\vartheta_i \in R_{\geq 0}$, and a function $\alpha_i \in \mathcal{K}_\infty$ such that*

1. *a) $\forall z_i^0 \in X_{0_i} \cap X_{S_i}$, $\exists \hat{z}_i^0 \in \hat{X}_{0_i} \cap \hat{X}_{S_i}$, s.t. $\mathcal{S}_i(z_i^0, \hat{z}_i^0) \leq \varepsilon_i$;*
   b) $\forall \hat{z}_i^0 \in \hat{X}_{0_i} \setminus \hat{X}_{S_i}$, $\exists z_i^0 \in X_{0_i} \setminus X_{S_i}$, s.t. $\mathcal{S}_i(z_i^0, \hat{z}_i^0) \leq \varepsilon_i$;
2. *$\forall z_i \in X_i, \forall \hat{z}_i \in \hat{X}_i$, $\alpha_i(\|\mathcal{H}_i(z_i) - \hat{\mathcal{H}}_i(\hat{z}_i)\|) \leq \mathcal{S}_i(z_i, \hat{z}_i)$;*
3. *$\forall z_i \in X_i, \forall \hat{z}_i \in \hat{X}_i$ s.t. $\mathcal{S}_i(z_i, \hat{z}_i) \leq \varepsilon_i$, $\forall w_i \in W_i$, $\forall \hat{w}_i \in \hat{W}_i$ s.t. $\|w_i - \hat{w}_i\| \leq \vartheta_i$, one has:*
   a) $\forall u_i \in U_i, \forall z_i^+ \in \mathcal{F}_i(z_i, u_i, w_i)$, $\exists \hat{u}_i \in \hat{U}_i$, $\exists \hat{z}_i^+ \in \hat{\mathcal{F}}_i(\hat{z}_i, \hat{u}_i, \hat{w}_i)$ s.t. $\mathcal{S}_i(z_i^+, \hat{z}_i^+) \leq \varepsilon_i$;
   b) $\forall \hat{u}_i \in \hat{U}_i, \forall \hat{z}_i^+ \in \hat{\mathcal{F}}_i(\hat{z}_i, \hat{u}_i, \hat{w}_i)$, $\exists u_i \in U_i$, $\exists z_i^+ \in \mathcal{F}_i(z_i, u_i, w_i)$ s.t. $\mathcal{S}_i(z_i^+, \hat{z}_i^+) \leq \varepsilon_i$.

Note that the local ε_i-InitSOPSFs are mainly proposed for constructing a ε-InitSOPSF for the networks and they are not directly used for deducing approximate initial-state opacity-preserving simulation relation. Similarly, we introduce a notion of local ε_i-CurSOPSFs for subsystems that can be used to establish ε-CurSOPSF for networks of switched systems.

Definition 6.3.16. *Consider transition systems* $T(\Sigma_i) = (X_i, X_{0_i}, X_{S_i}, U_i, W_i, \mathcal{F}_i, Y_i, \mathcal{H}_i)$ *and* $T(\hat{\Sigma}_i) = (\hat{X}_i, \hat{X}_{0_i}, \hat{X}_{S_i}, \hat{U}_i, \hat{W}_i, \hat{\mathcal{F}}_i, \hat{Y}_i, \hat{\mathcal{H}}_i)$*, for all* $i \in [1; N]$*, where* $\hat{W}_i \subseteq W_i$ *and* $\hat{Y}_i \subseteq Y_i$*. For* $\varepsilon_i \in \mathbb{R}_{\geq 0}$*, a function* $\mathcal{S}_i : X_i \times \hat{X}_i \to \mathbb{R}_{\geq 0}$ *is called a local* ε_i*-CurSOPSF from* $T(\Sigma_i)$ *to* $T(\hat{\Sigma}_i)$ *if there exist a constant* $\vartheta_i \in R_{\geq 0}$*, and a function* $\alpha_i \in \mathcal{K}_\infty$ *such that*

1. $\forall z_i^0 \in X_{0_i}$, $\exists \hat{z}_i^0 \in \hat{X}_{0_i}$, *s.t.* $\mathcal{S}_i(z_i^0, \hat{z}_i^0) \leq \varepsilon_i$;
2. $\forall z_i \in X_i, \forall \hat{z}_i \in \hat{X}_i$, $\alpha_i(\|\mathcal{H}_i(z_i) - \hat{\mathcal{H}}_i(\hat{z}_i)\|) \leq \mathcal{S}_i(z_i, \hat{z}_i)$;
3. $\forall z_i \in X_i, \forall \hat{z}_i \in \hat{X}_i$ *s.t.* $\mathcal{S}_i(z_i, \hat{z}_i) \leq \varepsilon_i$, $\forall w_i \in W_i$, $\forall \hat{w}_i \in \hat{W}_i$ *s.t.* $\|w_i - \hat{w}_i\| \leq \vartheta_i$, *one has:*
   a) $\forall u_i \in U_i$, $\forall z_i^+ \in \mathcal{F}_i(z_i, u_i, w_i)$, $\exists \hat{u}_i \in \hat{U}_i$, $\exists \hat{z}_i^+ \in \hat{\mathcal{F}}_i(\hat{z}_i, \hat{u}_i, \hat{w}_i)$ *s.t.* $\mathcal{S}_i(z_i^+, \hat{z}_i^+) \leq \varepsilon_i$;
   b) $\forall u_i \in U_i$, $\forall z_i^+ \in \mathcal{F}_i(z_i, u_i, w_i)$ *s.t.* $z_i^+ \in X_{S_i}$, $\exists \hat{u}_i \in \hat{U}_i$, $\exists \hat{z}_i^+ \in \hat{\mathcal{F}}_i(\hat{z}_i, \hat{u}_i, \hat{w}_i)$ *with* $\hat{z}_i^+ \in \hat{X}_{S_i}$ *s.t.* $\mathcal{S}_i(z_i^+, \hat{z}_i^+) \leq \varepsilon_i$;
   c) $\forall \hat{u}_i \in \hat{U}_i$, $\forall \hat{z}_i^+ \in \hat{\mathcal{F}}_i(\hat{z}_i, \hat{u}_i, \hat{w}_i)$, $\exists u_i \in U_i$, $\exists z_i^+ \in \mathcal{F}_i(z_i, u_i, w_i)$ *s.t.* $\mathcal{S}_i(z_i^+, \hat{z}_i^+) \leq \varepsilon_i$;
   d) $\forall \hat{u}_i \in \hat{U}_i$, $\forall \hat{z}_i^+ \in \hat{\mathcal{F}}_i(\hat{z}_i, \hat{u}_i, \hat{w}_i)$ *s.t.* $\hat{z}_i^+ \in \hat{X}_i \setminus \hat{X}_{S_i}$, $\exists u_i \in U_i$, $\exists z_i^+ \in \mathcal{F}_i(z_i, u_i, w_i)$ *with* $z_i^+ \in X_i \setminus X_{S_i}$ *s.t.* $\mathcal{S}_i(z_i^+, \hat{z}_i^+) \leq \varepsilon_i$.

We say that $T(\hat{\Sigma}_i)$ is an abstraction of $T(\Sigma_i)$ if there exists a local ε_i-InitSOPSF, or ε_i-CurSOPSF, from $T(\Sigma_i)$ to $T(\hat{\Sigma}_i)$. In addition, if $T(\hat{\Sigma}_i)$ is finite ($\hat{X}_i$ and $\hat{W}_i$ are finite sets), system $T(\hat{\Sigma}_i)$ is called a finite abstraction (symbolic model) of $T(\Sigma_i)$, and is denoted by $T(\Sigma_i) \preceq_L^{\varepsilon_i} T(\hat{\Sigma}_i)$.

Next, we show how to compose the above defined local simulation functions so that it can be used to quantify the distance between two networks in terms of preserving approximate opacity.

In this subsection, we provide one of the main results of the section. The following theorem provides a compositional approach for the construction of an opacity-preserving simulation function from $T(\Sigma)$ to $T(\hat{\Sigma})$ via the proposed local ε_i-InitSOPSF (resp. ε_i-CurSOPSF) from $T(\Sigma_i)$ to $T(\hat{\Sigma}_i)$.

Theorem 6.3.17. *Consider network* $T(\Sigma) = \mathcal{I}(T(\Sigma_1), \ldots, T(\Sigma_N))$*. Assume that each* $T(\Sigma_i)$ *admits an abstraction* $T(\hat{\Sigma}_i)$ *together with a local* ε_i*-InitSOPSF (resp. CurSOPSF)* $\mathcal{S}_i$*, associated with function* α_i *and constant* ϑ_i *as in Definition 6.3.15 (resp. Definition 6.3.16). Let* $\varepsilon = \max_i \varepsilon_i$*. If* $\forall i \in [1; N]$, $\forall j \in \mathcal{N}_i$,

$$\alpha_j^{-1}(\varepsilon_j) + \phi_{ij} \leq \vartheta_i, \tag{6.3.8}$$

where ϕ_{ij} *is an internal input quantization parameter for constructing the local finite abstractions* $T(\hat{\Sigma}_i)$*, then, function* $\mathcal{S} : X \times \hat{X} \to \mathbb{R}_{\geq 0}$ *defined as*

$$\mathcal{S}(z, \hat{z}) := \max_i \{\frac{\varepsilon}{\varepsilon_i} \mathcal{S}_i(z_i, \hat{z}_i)\}, \tag{6.3.9}$$

is an ε-InitSOPSF (resp. CurSOPSF) from $T(\Sigma) = \mathcal{I}(T(\Sigma_1), \ldots, T(\Sigma_N))$ to $T(\hat{\Sigma}) = \hat{\mathcal{I}}(T(\hat{\Sigma}_1), \ldots, T(\hat{\Sigma}_N))$.

Proof. First, we show that condition 1a) in Definition 6.3.11 holds. Consider any $z^0 \in X_0 \cap X_S$. For any system $T(\Sigma_i)$ and the corresponding ε_i-InitSOPSF $\mathcal{S}_i$, from the definition of $\mathcal{S}_i$, we have $\forall z_i^0 \in X_{0_i} \cap X_{S_i}$, $\exists \hat{z}_i^0 \in \hat{X}_{0_i} \cap \hat{X}_{s_i}$ s.t. $\mathcal{S}_i(z_i^0, \hat{z}_i^0) \leq \varepsilon_i$. Then, from the definition of $\mathcal{S}$ in (6.3.9) we get $\mathcal{S}(z^0, \hat{z}^0) \leq \varepsilon$, where $\hat{z}^0 \in \hat{X}_0 \cap \hat{X}_s$. Thus, condition 1(a) in Definition 6.3.11 holds. Condition 1b) can be proved in the same way, thus is omitted. Now, we show that condition 2 in Definition 6.3.11 holds for some $\mathcal{K}_\infty$ function α. Consider any $z = [z_1; \ldots; z_N] \in X$ and $\hat{z} = [\hat{z}_1; \ldots; \hat{z}_N] \in \hat{X}$. Then, using condition 2 in Definition 6.3.15, one gets

$$\begin{aligned}
&\|\mathcal{H}(z) - \hat{\mathcal{H}}(\hat{z})\| = \max_i\{\|\mathcal{H}_{ii}(z_i) - \hat{\mathcal{H}}_{ii}(\hat{z}_i)\|\} \\
&\leq \max_i\{\|\mathcal{H}_i(z_i) - \hat{\mathcal{H}}_i(\hat{z}_i)\|\} \leq \max_i\{\alpha_i^{-1} \circ \mathcal{S}_i(z_i, \hat{z}_i)\} \leq \hat{\alpha} \circ \max_i\{\frac{\varepsilon}{\varepsilon_i}\mathcal{S}_i(z_i, \hat{z}_i)\},
\end{aligned}$$

where $\hat{\alpha} = \max_i\{\alpha_i^{-1}\}$. By defining $\alpha = \hat{\alpha}^{-1}$, one obtains

$$\alpha(\|\mathcal{H}(z) - \hat{\mathcal{H}}(\hat{z})\|) \leq \mathcal{S}(z, \hat{z}),$$

which satisfies condition 2 in Definition 6.3.11. Now, we show that condition 3 holds. Let us consider any $z \in X$ and $\hat{z} \in \hat{X}$ such that $\mathcal{S}(z, \hat{z}) \leq \varepsilon$. It can be seen that from the structure of $\mathcal{S}$ in (6.3.9), we get $\mathcal{S}_i(z_i, \hat{z}_i) \leq \varepsilon_i$, $\forall i \in [1; N]$. For each pair of systems $T(\Sigma_i)$ and $T(\hat{\Sigma}_i)$, the internal inputs satisfy the chain of inequality

$$\begin{aligned}
&\|w_i - \hat{w}_i\| = \max_{j \in \mathcal{N}_i}\{\|w_{ij} - \hat{w}_{ij}\|\} = \max_{j \in \mathcal{N}_i}\{\|y_{ji} - \hat{y}_{ji} + \hat{y}_{ji} - \hat{w}_{ij}\|\} \\
&\leq \max_{j \in \mathcal{N}_i}\{\|y_{ji} - \hat{y}_{ji}\| + \phi_{ij}\} \leq \max_{j \in \mathcal{N}_i}\{\|\mathcal{H}_j(z_j) - \hat{\mathcal{H}}_j(\hat{z}_j)\| + \phi_{ij}\} \\
&\leq \max_{j \in \mathcal{N}_i}\{\alpha_j^{-1} \circ \mathcal{S}_j(z_j, \hat{z}_j) + \phi_{ij}\} \leq \max_{j \in \mathcal{N}_i}\{\alpha_j^{-1}(\varepsilon_j) + \phi_{ij}\}.
\end{aligned}$$

Using (6.3.8), one has $\|w_i - \hat{w}_i\| \leq \vartheta_i$. Therefore, by condition 3a) in Definition 6.3.15, for each pair of systems $T(\Sigma_i)$ and $T(\hat{\Sigma}_i)$, one has $\forall u_i \in U_i$, $\forall z_i^+ \in \mathcal{F}_i(z_i, u_i, w_i)$, there exists $\hat{u}_i \in \hat{U}_i$ and $\hat{z}_i^+ \in \hat{\mathcal{F}}_i(\hat{z}_i, \hat{u}_i, \hat{w}_i)$ such that $\mathcal{S}_i(z_i^+, \hat{z}_i^+) \leq \varepsilon_i$. As a result, we get $\forall u = [u_1; \ldots; u_N] \in U$, $\forall z^+ \in \mathcal{F}(z, u)$, there exists $\hat{u} = [\hat{u}_1; \ldots; \hat{u}_N] \in \hat{U}$ and $\hat{z}^+ \in \hat{\mathcal{F}}(\hat{z}, \hat{u})$ such that $\mathcal{S}(z^+, \hat{z}^+) = \max_i\{\frac{\varepsilon}{\varepsilon_i}\mathcal{S}_i(z_i^+, \hat{z}_i^+)\} \leq \varepsilon$. Therefore, condition 3a) in Definition 6.3.11 is satisfied with $\varepsilon = \max_i \varepsilon_i$. In addition, by condition 3b) in Definition 6.3.15, for each pair of systems $T(\Sigma_i)$ and $T(\hat{\Sigma}_i)$, one has $\forall \hat{u}_i \in \hat{U}_i$, $\forall \hat{z}_i^+ \in \hat{\mathcal{F}}_i(\hat{z}_i, \hat{u}_i, \hat{w}_i)$, there exists $u_i \in U_i$ and $z_i^+ \in \mathcal{F}_i(z_i, u_i, w_i)$ such that $\mathcal{S}_i(z_i^+, \hat{z}_i^+) \leq \varepsilon_i$. As a result, we get $\forall \hat{u} = [\hat{u}_1; \ldots; \hat{u}_N] \in \hat{U}$, $\forall \hat{z}^+ \in \hat{\mathcal{F}}(\hat{z}, \hat{u})$, there exists $u = [u_1; \ldots; u_N] \in U$ and $z^+ \in \mathcal{F}(z, u)$ such that $\mathcal{S}(z^+, \hat{z}^+) = \max_i\{\frac{\varepsilon}{\varepsilon_i}\mathcal{S}_i(z_i^+, \hat{z}_i^+)\} \leq \varepsilon$. It follows that condition 3b) in Definition 6.3.11 is satisfied as well. Therefore, we conclude that $\mathcal{S}$ is an ε-InitSOPSF from $T(\Sigma) = \mathcal{I}(T(\Sigma_1), \ldots, T(\Sigma_N))$ to $T(\hat{\Sigma}) = \hat{\mathcal{I}}(T(\hat{\Sigma}_1), \ldots, T(\hat{\Sigma}_N))$. Note that by following similar lines of reasoning as above, one can prove that $\mathcal{S}$ is also an ε-CurSOPSF from $T(\Sigma) = \mathcal{I}(T(\Sigma_1), \ldots, T(\Sigma_N))$ to $T(\hat{\Sigma}) = \hat{\mathcal{I}}(T(\hat{\Sigma}_1), \ldots, T(\hat{\Sigma}_N))$. □

Till here, we have seen that one can construct an abstraction of a network of switched systems by interconnecting local abstractions of the subsystems. The overall InitSOPSF (resp. CurSOPSF) between two networks is established by composing local InitSOPSFs (resp. local CurSOPSFs) as well. This abstract network allows us to check approximate opacity property over the simpler abstract network and carry the results back to the concrete network using the results provided in Corollary 6.3.14.

6.3.3 Compositionalisty Results

Next, we are going to impose certain conditions on the dynamics of the subsystems, such that one can construct proper abstractions for all of the subsystems together with the corresponding local InitSOPSFs or CurSOPSFs.

6.3.3.1 Construction of Local Finite Abstractions

In this section, we are going to explore how to construct finite abstractions together with local InitSOPSFs or CurSOPSFs for subsystems. The switched subsystems $\Sigma = (\mathbb{X}, \mathbb{X}_0, \mathbb{X}_S, P, W, F, Y, h)$ are assumed to be infinite and deterministic. Moreover, we assume the output map h satisfies the following general Lipschitz assumption: there exists an $\ell \in \mathcal{K}_\infty$ such that: $\|h(x) - h(y)\| \leq \ell(\|x - y\|)$ for all $x, y \in \mathbb{X}$. Here, we also use Σ_p to denote a switched subsystem Σ in (6.3.1) with constant switching signal $\mathsf{p}(k) = p,\ \forall k \in \mathbb{N}$.

Note that throughout this subsection, we are mainly talking about switched subsystems rather than the overall network. However, for the sake of better readability, we omit index i of subsystems throughout the text in this subsection, e.g., we write Σ and $T(\Sigma)$ instead of Σ_i and $T(\Sigma_i)$, respectively.

Here, we establish a local ε-InitSOPSF or ε-CurSOPSF between $T(\Sigma)$ and its finite abstraction by assuming that, for all $p \in P$, Σ_p is incrementally input-to-state stable (δ-ISS) (cf. Section 2.4). We restate the definition of incremental input-to-state stability for a switched subsystem as follows.

Definition 6.3.18. *A system Σ_p is δ-ISS if there exist a so-called δ-ISS Lyapunov function $V_p : X \times X \to \mathbb{R}_{\geq 0}$, $\underline{\alpha}_p, \overline{\alpha}_p, \rho_p \in \mathcal{K}_\infty$, and constant $0 < \kappa_p < 1$, such that for all $x, \hat{x} \in \mathbb{X}$, and for all $w, \hat{w} \in W$*

$$\underline{\alpha}_p(\|x - \hat{x}\|) \leq V_p(x, \hat{x}) \leq \overline{\alpha}_p(\|x - \hat{x}\|), \tag{6.3.10}$$

$$V_p(f_p(x, w), f_p(\hat{x}, \hat{w})) \leq \kappa_p V_p(x, \hat{x}) + \rho_p(\|w - \hat{w}\|). \tag{6.3.11}$$

Remark 6.3.19. *We say that V_p, $\forall p \in P$, are multiple δ-ISS Lyapunov functions for subsystem Σ if it satisfies* (6.3.10) *and* (6.3.11)*. Moreover, if $V_p = V_{p^+}, \forall p, p^+ \in P$, we omit the index p in* (6.3.10)*,* (6.3.11)*, and say that V is a common δ-ISS Lyapunov function for system Σ. We refer interested readers to [106] for more details on common and multiple Lyapunov functions for switched systems.*

Next, we provide an approach, inspired by [53], to construct a local finite abstraction $T(\hat{\Sigma})$ of transition system $T(\Sigma)$ associated to the switched subsystem Σ in which each mode Σ_p is δ-ISS.

Definition 6.3.20. *Consider a transition system $T(\Sigma) = (X, X_0, X_S, U, W, \mathcal{F}, Y, \mathcal{H})$, associated to the switched subsystem $\Sigma = (\mathbb{X}, \mathbb{X}_0, \mathbb{X}_S, P, W, F, Y, h)$, where $\mathbb{X}, W$ are assumed to be finite unions of boxes. Let Σ_p be δ-ISS as in Definition 6.3.18. Then one can construct a finite abstraction $T(\hat{\Sigma}) = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{W}, \hat{\mathcal{F}}, \hat{Y}, \hat{\mathcal{H}})$ where:*

- $\hat{X} = \hat{\mathbb{X}} \times P \times \{0, \ldots, k_d - 1\}$, *where* $\hat{\mathbb{X}} = [\mathbb{X}]_\eta$ *and* $0 < \eta \leq \min\{span(\mathbb{X}_S), span(\mathbb{X} \setminus \mathbb{X}_S)\}$ *is the state set quantization parameter;*
- $\hat{X}_0 = \hat{\mathbb{X}}_0 \times P \times \{0\}$, *where* $\hat{\mathbb{X}}_0 = [\mathbb{X}_0]_\eta$*;*
- $\hat{X}_S = \hat{\mathbb{X}}_S \times P \times \{0, \cdots, k_d - 1\}$, *where* $\hat{\mathbb{X}}_S = [\mathbb{X}_S^\theta]_\eta$, *and* $\mathbb{X}_S^\theta = \{x \in \mathbb{X} \mid \exists \bar{x} \in \mathbb{X}_S, \|x - \bar{x}\| \leq \theta\}$ *denotes the θ-expansion of set $\mathbb{X}_S$ where $\theta > 0$ is a design parameter;*
- $\hat{U} = U = P$*;*
- $(\hat{x}^+, p^+, l^+) \in \hat{\mathcal{F}}((\hat{x}, p, l), \hat{u}, \hat{w})$ *if and only if* $\|f_p(\hat{x}, \hat{w}) - \hat{x}^+\| \leq \eta$, $\hat{u} = p$ *and one of the following scenarios hold:*
  - $l < k_d - 1$, $p^+ = p$ *and* $l^+ = l + 1$*;*
  - $l = k_d - 1$, $p^+ = p$ *and* $l^+ = k_d - 1$*;*
  - $l = k_d - 1$, $p^+ \neq p$ *and* $l^+ = 0$*;*
- $\hat{Y} = \{\mathcal{H}(\hat{x}, p, l) | (\hat{x}, p, l) \in \hat{X}\}$*;*
- $\hat{\mathcal{H}} : \hat{X} \to \hat{Y}$, *defined as* $\hat{\mathcal{H}}(\hat{x}, p, l) = \mathcal{H}(\hat{x}, p, l) = h(\hat{x})$*;*
- $\hat{W} = [W]_\phi$, *where ϕ, satisfying* $0 < \|\phi\| \leq span(W)$, *is the internal input set quantization parameter.*

Note that in the case when the concrete switched subsystem Σ admits a common δ-ISS Lyapunov function as in Remark 6.3.19, Definition 6.3.20 boils down to the following.

Definition 6.3.21. *Consider a transition system $T(\Sigma) = (X, X_0, X_S, U, W, \mathcal{F}, Y, \mathcal{H})$, associated to the switched subsystem $\Sigma = (\mathbb{X}, \mathbb{X}_0, \mathbb{X}_S, P, W, F, Y, h)$, where $\mathbb{X}, W$ are assumed to be finite unions of boxes. Suppose Σ admits a common δ-ISS Lyapunov function as in Remark 6.3.19. Then one can construct a finite abstraction $T(\hat{\Sigma}) = (\hat{X}, \hat{X}_0, \hat{X}_S, \hat{U}, \hat{W}, \hat{\mathcal{F}}, \hat{Y}, \hat{\mathcal{H}})$ where:*

- $\hat{X} = [\mathbb{X}]_\eta$, *where* $0 < \eta \leq \min\{span(\mathbb{X}_S), span(\mathbb{X} \setminus \mathbb{X}_S)\}$ *is the state set quantization parameter;*
- $\hat{X}_0 = [\mathbb{X}_0]_\eta$*;*
- $\hat{X}_S = [\mathbb{X}_S^\theta]_\eta$, *where* $\mathbb{X}_S^\theta = \{x \in \mathbb{X} \mid \exists \bar{x} \in \mathbb{X}_S, \|x - \bar{x}\| \leq \theta\}$ *denotes the θ-expansion of set $\mathbb{X}_S$ where $\theta > 0$ is a design parameter;*
- $\hat{U} = P$*;*

- $\hat{x}^+ \in \hat{\mathcal{F}}(\hat{x}, \hat{u}, \hat{w})$ *if and only if* $\|f_{\hat{u}}(\hat{x}, \hat{w}) - \hat{x}^+\| \leq \eta$;
- $\hat{Y} = \{h(\hat{x}) | \hat{x} \in \hat{X}\}$;
- $\hat{\mathcal{H}} : \hat{X} \to \hat{Y}$, *defined as* $\hat{\mathcal{H}}(\hat{x}) = h(\hat{x})$;
- $\hat{W} = [W]_\phi$, *where* ϕ, *satisfying* $0 < \|\phi\| \leq span(W)$, *is the internal input set quantization parameter.*

In order to construct a local ε-InitSOPSF or ε-CurSOPSF from $T(\Sigma)$ to $T(\hat{\Sigma})$, we raise the following assumptions on functions V_p appeared in Definition 6.3.18, which are used to prove some of the main results later.

Assumption 6.3.22. *There exists* $\mu \geq 1$ *such that*

$$\forall x, y \in \mathbb{X}, \quad \forall p, q \in P, \quad V_p(x, y) \leq \mu V_q(x, y). \tag{6.3.12}$$

Assumption 6.3.22 is an incremental version of a similar assumption in [195] that is used to prove input-to-state stability of switched systems under constrained switching signals.

Assumption 6.3.23. *For all* $p \in P$, *there exists a* $\mathcal{K}_\infty$ *function* γ_p *such that*

$$\forall x, y, z \in \mathbb{X}, \quad V_p(x, y) \leq V_p(x, z) + \gamma_p(\|y - z\|). \tag{6.3.13}$$

Assumption 6.3.23 is non-restrictive as shown in [216] provided that one is interested to work on a compact subset of $\mathbb{X}$.

Now, we establish the relation between $T(\Sigma)$ and $T(\hat{\Sigma})$, introduced above, via the notion of local ε-InitSOPSF as in Definition 6.3.15.

Theorem 6.3.24. *Consider a switched subsystems* $\Sigma = (\mathbb{X}, \mathbb{X}_0, \mathbb{X}_S, P, W, F, Y, h)$ *with its equivalent transition system* $T(\Sigma) = (X, X_0, X_S, U, W, \mathcal{F}, Y, \mathcal{H})$. *Suppose* Σ_p *is* δ*-ISS as in Definition 6.3.18, with a function* V_p *equipped with functions* $\underline{\alpha}_p, \overline{\alpha}_p, \rho_p$ *and constant* κ_p, *and Assumptions 6.3.22 and 6.3.23 hold. Let* $\epsilon > 1$. *For any design parameters* $\varepsilon, \vartheta \in R_{\geq 0}$, *let* $T(\hat{\Sigma})$ *be a finite abstraction of* $T(\Sigma)$ *constructed as in Definition 6.3.20 with any quantization parameter* $\eta \in R_{>0}$ *satisfying*

$$\eta \leq \min\{\hat{\gamma}^{-1}((1 - \kappa)\varepsilon - \rho(\vartheta)), \overline{\alpha}^{-1}(\varepsilon)\}, \tag{6.3.14}$$

where $\kappa = \max_{p \in P} \left\{ \kappa_p^{\frac{\epsilon - 1}{\epsilon}} \right\}$, $\rho = \max_{p \in P} \left\{ \kappa_p^{-\frac{k_d}{\epsilon}} \rho_p \right\}$, $\hat{\gamma} = \max_{p \in P} \left\{ \kappa_p^{-\frac{k_d}{\epsilon}} \gamma_p \right\}$, $\overline{\alpha} = \max_{p \in P} \left\{ \kappa_p^{-\frac{l}{\epsilon}} \overline{\alpha}_p \right\}$. *If,* $\forall p \in P$, $k_d \geq \epsilon \frac{\ln(\mu)}{\ln(\frac{1}{\kappa_p})} + 1$, *then function* $\mathcal{V}$ *defined as*

$$\mathcal{V}((x, p, l), (\hat{x}, p, l)) := V_p(x, \hat{x}) \kappa_p^{\frac{-l}{\epsilon}}, \tag{6.3.15}$$

is a local ε*-InitSOPSF from* $T(\Sigma)$ *to* $T(\hat{\Sigma})$ *and from* $T(\hat{\Sigma})$ *to* $T(\Sigma)$.

Proof. We start by proving condition 1 in Definition 6.3.15. Consider any initial and secret state $(x^0, p^0, 0) \in X_0 \cap X_S$ in $T(\Sigma)$. From Definition 6.3.20, for every $(x^0, p^0, 0) \in X_0 \cap X_S$, there always exists $(\hat{x}^0, p^0, 0) \in \hat{X}_0 \cap \hat{X}_S$ such that $\|x^0 - \hat{x}^0\| \leq \eta$. Hence, using (6.3.10), there exists $(\hat{x}^0, p^0, 0) \in \hat{X}_0 \cap \hat{X}_S$ with $\mathcal{V}((x^0, p^0, 0), (\hat{x}^0, p^0, 0)) \leq \frac{\overline{\alpha}_p(\|x^0 - \hat{x}^0\|)}{\kappa_p^{\frac{l}{\epsilon}}} \leq \frac{\overline{\alpha}_p(\eta)}{\kappa_p^{\frac{l}{\epsilon}}}$, and condition 1(a) is satisfied with $\overline{\alpha} = \max_{p \in P} \left\{ \kappa_p^{-\frac{l}{\epsilon}} \overline{\alpha}_p \right\}$ and $\overline{\alpha}(\eta) \leq \varepsilon$ by (6.3.14). For every $(\hat{x}^0, p^0, 0) \in \hat{X}_0 \setminus \hat{X}_S$, by choosing $x^0 = \hat{x}^0$ with $(x^0, p^0, 0)$ also being inside $X_0 \setminus X_S$, we get $\mathcal{V}((x^0, p^0, 0), (\hat{x}^0, p^0, 0)) = 0 \leq \varepsilon$. Hence, condition 1(b) in Definition 6.3.15 holds as well.

Next, we show condition 2 in Definition 6.3.15 holds. Given the Lipschitz assumption on h and since, $\forall p \in P$, Σ_p is δ-ISS, from (6.3.10), $\forall (x, p, l) \in X$ and $\forall (\hat{x}, p, l) \in \hat{X}$, we have

$$\begin{aligned}
&\|\mathcal{H}(x, p, l) - \hat{\mathcal{H}}(\hat{x}, p, l)\| = \|h(x) - \hat{h}(\hat{x})\| \leq \ell(\|x - \hat{x}\|) \\
&\leq \ell \circ \underline{\alpha}_p^{-1}(V_p(x, \hat{x})) = \ell \circ \underline{\alpha}_p^{-1} \left(\kappa_p^{\frac{l}{\epsilon}} \mathcal{V}((x, p, l), (\hat{x}, p, l)) \right) \\
&\leq \ell \circ \underline{\alpha}_p^{-1} \left(\mathcal{V}((x, p, l), (\hat{x}, p, l)) \right) \leq \hat{\alpha} \left(\mathcal{V}((x, p, l), (\hat{x}, p, l)) \right),
\end{aligned}$$

where $\hat{\alpha} = \max_{p \in P} \{ \ell \circ \underline{\alpha}_p^{-1} \}$. By defining $\alpha = \hat{\alpha}^{-1}$, one obtains

$$\alpha(\|\mathcal{H}(x, p, l) - \hat{\mathcal{H}}(\hat{x}, p, l)\|) \leq \mathcal{V}((x, p, l), (\hat{x}, p, l)),$$

satisfying condition 2. Now we show condition 3 in Definition 6.3.15. From (6.3.13), $\forall x \in \mathbb{X}, \forall \hat{x} \in \hat{\mathbb{X}}$, $\forall w \in W, \forall \hat{w} \in \hat{W}$, we have

$$V_p(f_p(x, w), \hat{x}^+) \leq V_p(f_p(x, w), f_p(\hat{x}, \hat{w})) + \gamma_p(\|\hat{x}^+ - f_p(\hat{x}, \hat{w})\|),$$

for any $\hat{x}^+$ such that $(\hat{x}^+, p^+, l^+) \in \hat{\mathcal{F}}((\hat{x}, p, l), \hat{u}, \hat{w})$. Now, from Definition 6.3.20, the above inequality reduces to

$$V_p(f_p(x, w), \hat{x}^+) \leq V_p(f_p(x, w), f_p(\hat{x}, \hat{w})) + \gamma_p(\eta).$$

Note that by (6.3.11), one gets

$$V_p(f_p(x, w), f_p(\hat{x}, \hat{w})) \leq \kappa_p V_p(x, \hat{x}) + \rho_p(\|w - \hat{w}\|).$$

Hence, $\forall x \in \mathbb{X}, \forall \hat{x} \in \hat{\mathbb{X}}$, $\forall w \in W, \forall \hat{w} \in \hat{W}$, one obtains

$$V_p(f_p(x, w), \hat{x}^+) \leq \kappa_p V_p(x, \hat{x}) + \rho_p(\|w - \hat{w}\|) + \gamma_p(\eta), \tag{6.3.16}$$

for any $\hat{x}^+$ such that $(\hat{x}^+, p^+, l^+) \in \hat{\mathcal{F}}((\hat{x}, p, l), \hat{u}, \hat{w})$. Now, in order to show function $\mathcal{V}$ defined in (6.3.15) satisfies condition 3 in Definition 6.3.15, we consider the different scenarios in Definition 6.3.20:

- $l < k_d - 1$, $p^+ = p$ and $l^+ = l + 1$, using (6.3.16) and $k_d > l + 1$, we have

$$\begin{aligned}\mathcal{V}((x^+,p^+,l^+),(\hat{x}^+,p^+,l^+)) &= \frac{V_{p^+}(x^+,\hat{x}^+)}{\kappa_p^{\frac{l^+}{\epsilon}}} = \frac{V_p(f_p(x,w),\hat{x}^+)}{\kappa_p^{\frac{l+1}{\epsilon}}}\\ &\leq \frac{\kappa_p V_p(x,\hat{x}) + \rho_p(\|w-\hat{w}\|) + \gamma_p(\eta)}{\kappa_p^{\frac{l+1}{\epsilon}}} = \frac{\kappa_p}{\kappa_p^{\frac{1}{\epsilon}}}\frac{V_p(x,\hat{x})}{\kappa_p^{\frac{l}{\epsilon}}} + \frac{\rho_p(\|w-\hat{w}\|) + \gamma_p(\eta)}{\kappa_p^{\frac{l+1}{\epsilon}}}\\ &\leq \kappa_p^{\frac{\epsilon-1}{\epsilon}}\mathcal{V}((x,p,l),(\hat{x},p,l)) + \frac{\rho_p(\|w-\hat{w}\|) + \gamma_p(\eta)}{\kappa_p^{\frac{k_d}{\epsilon}}}.\end{aligned}$$

- $l = k_d - 1$, $p^+ = p$ and $l^+ = k_d - 1$, using (6.3.16) and $\frac{\epsilon-1}{\epsilon} < 1$, one gets

$$\begin{aligned}\mathcal{V}((x^+,p^+,l^+),(\hat{x}^+,p^+,l^+)) &= \frac{V_{p^+}(x^+,\hat{x}^+)}{\kappa_p^{\frac{l^+}{\epsilon}}} = \frac{V_p(f_p(x,w),\hat{x}^+)}{\kappa_p^{\frac{l}{\epsilon}}}\\ &\leq \frac{\kappa_p V_p(x,\hat{x}) + \rho_p(\|w-\hat{w}\|) + \gamma_p(\eta)}{\kappa_p^{\frac{l}{\epsilon}}} = \kappa_p\frac{V_p(x,\hat{x})}{\kappa_p^{\frac{l}{\epsilon}}} + \frac{\rho_p(\|w-\hat{w}\|) + \gamma_p(\eta)}{\kappa_p^{\frac{l}{\epsilon}}}\\ &\leq \kappa_p^{\frac{\epsilon-1}{\epsilon}}\mathcal{V}((x,p,l),(\hat{x},p,l)) + \frac{\rho_p(\|w-\hat{w}\|) + \gamma_p(\eta)}{\kappa_p^{\frac{k_d}{\epsilon}}}.\end{aligned}$$

- $l = k_d - 1$, $p^+ \neq p$ and $l^+ = 0$, using (6.3.16), $\mu\kappa_p^{\frac{k_d-1}{\epsilon}} \leq 1$, and $\frac{\epsilon-1}{\epsilon} < 1$, one has

$$\begin{aligned}\mathcal{V}((x^+,p^+,l^+),(\hat{x}^+,p^+,l^+)) &= \frac{V_{p^+}(x^+,\hat{x}^+)}{\kappa_{p^+}^{\frac{l^+}{\epsilon}}} \leq \mu V_p(f_p(x,w),\hat{x}^+)\\ &\leq \frac{\mu\kappa_p^{\frac{k_d-1}{\epsilon}}(\kappa_p V_p(x,\hat{x}) + \rho_p(\|w-\hat{w}\|) + \gamma_p(\eta))}{\kappa_p^{\frac{k_d-1}{\epsilon}}}\\ &= \kappa_p\frac{V_p(x,\hat{x})}{\kappa_p^{\frac{l}{\epsilon}}} + \frac{\rho_p(\|w-\hat{w}\|) + \gamma_p(\eta)}{\kappa_p^{\frac{l}{\epsilon}}}\\ &\leq \kappa_p^{\frac{\epsilon-1}{\epsilon}}\mathcal{V}((x,p,l),(\hat{x},p,l)) + \frac{\rho_p(\|w-\hat{w}\|) + \gamma_p(\eta)}{\kappa_p^{\frac{k_d}{\epsilon}}}.\end{aligned}$$

Note that $\forall p \in P, \mu\kappa_p^{\frac{k_d-1}{\epsilon}} \leq 1$, since $\forall p \in P, k_d \geq \epsilon\frac{\ln(\mu)}{\ln(\frac{1}{\kappa_p})} + 1$. Hence, $\forall (x,p,l) \in X$, $\forall (\hat{x},p,l) \in \hat{X}$, $\forall w \in W$, $\forall \hat{w} \in \hat{W}$, one gets

$$\mathcal{V}((x^+,p^+,l^+),(\hat{x}^+,p^+,l^+)) \leq \kappa\mathcal{V}((x,p,l),(\hat{x},p,l)) + \rho(\|w-\hat{w}\|) + \hat{\gamma}(\eta). \quad (6.3.17)$$

Now, we show the condition 3a) in Definition 6.3.15 holds. Let us consider any pair of states $(x,p,l) \in X$, $(\hat{x},p,l) \in \hat{X}$, satisfying $\mathcal{V}((x,p,l),(\hat{x},p,l)) \leq \varepsilon$, and any $w \in W$, $\hat{w} \in \hat{W}$ such that $\|w-\hat{w}\| \leq \vartheta$. Combining (6.3.17) with (6.3.14) for any $(x^+,p^+,l^+) \in \mathcal{F}((x,p,l),u,w)$ and any $(\hat{x}^+,p^+,l^+) \in \hat{\mathcal{F}}((\hat{x},p,l),\hat{u},\hat{w})$ with $\hat{u} = u$, one obtains:

$$\mathcal{V}((x^+,p^+,l^+),(\hat{x}^+,p^+,l^+)) \leq \kappa\varepsilon + \rho(\vartheta) + \hat{\gamma}(\hat{\gamma}^{-1}((1-\kappa)\varepsilon - \rho(\vartheta))) = \varepsilon, \quad (6.3.18)$$

which shows that condition 3a) is satisfied. Similarly, for any $(\hat{x}^+, p^+, l^+) \in \hat{\mathcal{F}}((\hat{x}, p, l), \hat{u}, \hat{w})$, condition 3b) is also satisfied using the same reasoning with $(x^+, p^+, l^+) \in \mathcal{F}((x, p, l), \hat{u}, w)$. Therefore, we conclude that $\mathcal{V}$ is a local ε-InitSOPSF from $T(\Sigma)$ to $T(\hat{\Sigma})$. Similarly, one can show that $\mathcal{V}$ is a also a local ε-InitSOPSF from $T(\hat{\Sigma})$ to $T(\Sigma)$. □

Note that a similar framework for constructing symbolic models of switched systems was first proposed in [53], where the results take a monolithic view of the concrete switched systems without considering the distinction between internal and external inputs and outputs. However, their distinction plays an important role in our proposed compositional scheme which allows us to build symbolic models for switched subsystems individually and then construct a symbolic model for the overall network by interconnecting those local ones.

Next, we provide a similar result as in Theorem 6.3.24, but tailored to approximate current-state opacity.

Theorem 6.3.25. *Consider a switched subsystems $\Sigma = (\mathbb{X}, \mathbb{X}_0, \mathbb{X}_S, P, W, F, Y, h)$ with its equivalent transition system $T(\Sigma) = (X, X_0, X_S, U, W, \mathcal{F}, Y, \mathcal{H})$. Suppose Σ_p is δ-ISS as in Definition 6.3.18, with a function V_p equipped with functions $\underline{\alpha}_p, \overline{\alpha}_p, \rho_p$ and constant κ_p, and Assumptions 6.3.22 and 6.3.23 hold. Let $\epsilon > 1$. For any design parameters $\varepsilon, \vartheta \in R_{\geq 0}$, let $T(\hat{\Sigma})$ be a finite abstraction of $T(\Sigma)$ constructed as in Definition 6.3.20 with any quantization parameters $\eta \in R_{>0}$ and $\theta \in R_{>0}$ satisfying*

$$\eta \leq \min\{\hat{\gamma}^{-1}((1-\kappa)\varepsilon - \rho(\vartheta)), \overline{\alpha}^{-1}(\varepsilon)\}; \tag{6.3.19}$$

$$\underline{\alpha}^{-1}(\varepsilon) \leq \theta, \tag{6.3.20}$$

where $\kappa = \max_{p\in P}\left\{\kappa_p^{\frac{\epsilon-1}{\epsilon}}\right\}$, $\rho = \max_{p\in P}\left\{\kappa_p^{-\frac{k_d}{\epsilon}}\rho_p\right\}$, $\hat{\gamma} = \max_{p\in P}\left\{\kappa_p^{-\frac{k_d}{\epsilon}}\gamma_p\right\}$, $\overline{\alpha} = \max_{p\in P}\left\{\kappa_p^{-\frac{l}{\epsilon}}\overline{\alpha}_p\right\}$, $\underline{\alpha} = \min_{p\in P}\left\{\kappa_p^{-\frac{l}{\epsilon}}\underline{\alpha}_p\right\}$. *If,* $\forall p \in P$, $k_d \geq \epsilon\frac{\ln(\mu)}{\ln(\frac{1}{\kappa_p})} + 1$, *then function $\mathcal{V}$ defined as*

$$\mathcal{V}((x, p, l), (\hat{x}, p, l)) := V_p(x, \hat{x})\kappa_p^{\frac{-l}{\epsilon}}, \tag{6.3.21}$$

is a local ε-CurSOPSF from $T(\Sigma)$ to $T(\hat{\Sigma})$.

Proof. We start by proving condition 1 in Definition 6.3.16. Consider any initial state $(x^0, p^0, 0) \in X_0$ in $T(\Sigma)$. Note that from Definition 6.3.20, we have $\hat{X}_0 = \hat{X}_0 \times P \times \{0\}$, where $\hat{X}_0 = [\mathbb{X}_0]_\eta$. Therefore, for every $(x^0, p^0, 0) \in X_0$, there always exists $(\hat{x}^0, p^0, 0) \in \hat{X}_0$ such that $\|x^0 - \hat{x}^0\| \leq \eta$. Hence, one gets $\mathcal{V}((x^0, p^0, 0), (\hat{x}^0, p^0, 0)) \leq \frac{\overline{\alpha}_p(\|x^0 - \hat{x}^0\|)}{\kappa_p^{\frac{l}{\epsilon}}} \leq \frac{\overline{\alpha}_p(\eta)}{\kappa_p^{\frac{l}{\epsilon}}}$ by (6.3.10), and condition 1 is satisfied with $\overline{\alpha} = \max_{p\in P}\left\{\kappa_p^{-\frac{l}{\epsilon}}\overline{\alpha}_p\right\}$ and $\overline{\alpha}(\eta) \leq \varepsilon$ by

(6.3.19). The proof for conditions 2, 3a), and 3c) in Definition 6.3.16 is similar to that of Theorem 6.3.24, and is omitted here.

For condition 3b), let us consider any $u \in U$ s.t. $(x^+, p^+, l^+) \in X_s$. By choosing $\hat{u} = u$ and following same reasoning as in Theorem 6.3.24, we obtain

$$\mathcal{V}((x^+, p^+, l^+), (\hat{x}^+, p^+, l^+)) \leq \varepsilon.$$

Additionally, by combining (6.3.10) and (6.3.21), one gets

$$\|x^+ - \hat{x}^+\| \overset{(6.3.10)}{\leq} \underline{\alpha}_p^{-1}(V_p(x^+, \hat{x}^+)) \overset{(6.3.21)}{=} \underline{\alpha}_p^{-1}\kappa_p^{\frac{l}{\epsilon}}(\mathcal{V}((x^+, p^+, l^+), (\hat{x}^+, p^+, l^+))) \leq \underline{\alpha}^{-1}(\varepsilon),$$

where $\underline{\alpha} = \min_{p \in P}\left\{\kappa_p^{-\frac{l}{\epsilon}}\underline{\alpha}_p\right\}$. Moreover, by (6.3.20), one gets $\|x^+ - \hat{x}^+\| \leq \underline{\alpha}^{-1}(\varepsilon) \leq \theta$. Note that by the structure of the abstraction as in Definition 6.3.20, we have $\hat{X}_S = \hat{X}_S \times P \times \{0, \cdots, k_d - 1\}$ where $\hat{X}_S = [X_S^\theta]_\eta$ and $X_S^\theta = \{x \in X \mid \exists \bar{x} \in X_S, \|x - \bar{x}\| \leq \theta\}$. This implies that $(\hat{x}^+, p^+, l^+) \in \hat{X}_s$, and thus, condition 3b) is satisfied as well. Condition 3d) of Definition 6.3.16 can be proved in a similar way and is omitted here. Therefore, we conclude that $\mathcal{V}$ is a local ε-CurSOPSF from $T(\Sigma)$ to $T(\hat{\Sigma})$. □

Remark 6.3.26. *If Σ admits a common δ-ISS Lyapunov function satisfying Assumption 6.3.23, then functions $\mathcal{V}$ defined in Theorems 6.3.24 and 6.3.25 reduce to $\mathcal{V}((x, p, l), (\hat{x}, p, l)) := V(x, \hat{x})$.*

Given the results of Theorems 6.3.17 and 6.3.24 (resp. 6.3.25), one can see that conditions (6.3.8) and (6.3.14) (resp. (6.3.19)) may not hold at the same time. In the following subsection, we will discuss about the inherent property that the network should have such that one can design suitable quantization parameters to satisfy conditions (6.3.8) and (6.3.14) (resp. (6.3.19)) simultaneously.

6.3.3.2 Compositional Construction of Opacity-Preserving Finite Abstractions

We raise the following assumption which provides a small-gain type condition, inspired by [39, Theorem 5.2], so that one can verify whether the competing conditions (6.3.8) and (6.3.14) (resp. (6.3.19)) can be satisfied simultaneously.

Assumption 6.3.27. *Consider network $\mathcal{I}(T(\Sigma_1), \ldots, T(\Sigma_N))$ induced by $N \in \mathbb{N}_{\geq 1}$ transition systems $T(\Sigma_i)$. Assume that each $T(\Sigma_i)$ and its finite abstraction $T(\hat{\Sigma}_i)$ admit a local ε_i-InitSOPSF (resp. ε_i-CurSOPSF) $\mathcal{V}_i$ defined in* (6.3.15) *(resp.* (6.3.21)*), associated with functions and constants κ_i, α_i, and ρ_i that appeared in Theorem 6.3.24 (resp. Theorem 6.3.25). Define*

$$\gamma_{ij} := \begin{cases} (1 - \kappa_i)^{-1}\rho_i \circ \alpha_j^{-1} & \text{if } j \in \mathcal{N}_i, \\ 0 & \text{otherwise}, \end{cases} \tag{6.3.22}$$

for all $i, j \in [1; N]$, and assume that functions γ_{ij} defined in (6.3.22) *satisfy*

$$\gamma_{i_1 i_2} \circ \gamma_{i_2 i_3} \circ \cdots \circ \gamma_{i_{r-1} i_r} \circ \gamma_{i_r i_1} < \mathrm{id}, \tag{6.3.23}$$

$\forall (i_1, \ldots, i_r) \in \{1, \ldots, N\}^r$, where $r \in \{1, \ldots, N\}$.

Algorithm 2: Compositional design of local quantization parameters $\eta_i \in \mathbb{R}_{>0}$ and $\phi_{ij} \in \mathbb{R}_{>0}$, $\forall i \in [1;N]$.

Input: The desired precision $\varepsilon \in \mathbb{R}_{>0}$; the simulation functions $\mathcal{V}_i$ equipped with functions κ_i, α_i, ρ_i, $\hat{\gamma}_i$, and $\overline{\alpha}_i$, $\forall i \in [1;N]$; functions σ_i, $\forall i \in [1;N]$, satisfying (6.3.24).

1 Choose $r \in \mathbb{R}_{>0}$ s.t. $\max\limits_{i\in[1;N]}\{\sigma_i(r)\} = \varepsilon$;

2 Set $\varepsilon_i = \sigma_i(r)$, $\forall i \in [1;N]$;

3 Design $\phi_{ij} \in \mathbb{R}_{>0}$ s.t. $\max\limits_{j\in\mathcal{N}_i}\{\phi_{ij}\} < \rho_i^{-1}((1-\kappa_i)\varepsilon_i) - \max\limits_{j\in\mathcal{N}_i}\{\alpha_j^{-1}(\varepsilon_j)\}, \forall i,j \in [1;N]$;

4 Set $\vartheta_i = \max\limits_{j\in\mathcal{N}_i}\{\alpha_j^{-1}(\varepsilon_j) + \phi_{ij}\}$, $\forall i \in [1;N]$;

5 Design $\eta_i \in \mathbb{R}_{>0}$ s.t. $\eta_i \leq \min\{\hat{\gamma}_i^{-1}((1-\kappa_i)\varepsilon_i - \rho_i(\vartheta_i)), \overline{\alpha}_i^{-1}(\varepsilon_i)\}$;

Output: Quantization parameters $\eta_i \in \mathbb{R}_{>0}$ and $\phi_{ij} \in \mathbb{R}_{>0}$, $\forall i \in [1;N]$.

Now, we show that, under the above small-gain assumption, one can always compositionally design local quantization parameters to satisfy conditions (6.3.8) and (6.3.14) (resp. (6.3.19)) simultaneously.

Theorem 6.3.28. *Suppose that Assumption 6.3.27 holds. Then, there always exist local quantization parameters η_i and ϕ_{ij}, $\forall i,j \in [1;N]$, as designed in Algorithm 2, such that* (6.3.8) *and* (6.3.14) *(resp.* (6.3.19)*) can be satisfied simultaneously.*

Proof. First, let us note that the small-gain condition (6.3.23) implies that $\exists \sigma_i \in \mathcal{K}_\infty$ satisfying $\forall i \in [1;N]$,

$$\max_{j\in\mathcal{N}_i}\{\gamma_{ij} \circ \sigma_j\} < \sigma_i, \tag{6.3.24}$$

see [39, Theorem 5.2]. Then, from (6.3.22), we have $\forall i \in [1;N]$,

$$\begin{aligned}&\max_{j\in\mathcal{N}_i}\{\gamma_{ij} \circ \sigma_j\} < \sigma_i \Longrightarrow \max_{j\in\mathcal{N}_i}\{(1-\kappa_i)^{-1}\rho_i \circ \alpha_j^{-1} \circ \sigma_j\} < \sigma_i\\ &\Longrightarrow \rho_i \circ \max_{j\in\mathcal{N}_i}\{\alpha_j^{-1} \circ \sigma_j\} < (1-\kappa_i)\sigma_i.\end{aligned} \tag{6.3.25}$$

Next, suppose that we are given a sequence of functions $\sigma_i \in \mathcal{K}_\infty$, $\forall i \in [1;N]$, satisfying (6.3.24). Assume we are given any desired precision ε as in Definition 6.3.11. Let us set $\varepsilon_i = \sigma_i(r)$, $\forall i \in [1;N]$, where $r \in \mathbb{R}_{>0}$ is chosen such that $\max\limits_i\{\sigma_i(r)\} = \varepsilon$. Then, we choose internal input quantization parameters ϕ_{ij}, $\forall i,j \in [1;N]$, such that

$$\max_{j\in\mathcal{N}_i}\{\phi_{ij}\} < \rho_i^{-1}((1-\kappa_i)\varepsilon_i) - \max_{j\in\mathcal{N}_i}\{\alpha_j^{-1}(\varepsilon_j)\}. \tag{6.3.26}$$

Now, by setting $\vartheta_i = \max_{j\in\mathcal{N}_i}\{\alpha_j^{-1}(\varepsilon_j) + \phi_{ij}\}$, and combining (6.3.25) and (6.3.26), one has $\forall i \in [1; N]$

$$\begin{aligned}\rho_i(\vartheta_i) &= \rho_i(\max_{j\in\mathcal{N}_i}\{\alpha_j^{-1}(\varepsilon_j) + \phi_{ij}\}) \\ &\leq \rho_i(\max_{j\in\mathcal{N}_i}\{\alpha_j^{-1}(\varepsilon_j) + \max_{j\in\mathcal{N}_i}\{\phi_{ij}\}) < (1-\kappa_i)\varepsilon_i. \end{aligned} \tag{6.3.27}$$

Thus, by (6.3.27), given any pair of parameters $(\varepsilon_i, \vartheta_i)$, one can always find suitable local parameters η_i to satisfy (6.3.14) (resp. (6.3.19)). Additionally, the selection of $\vartheta_i = \max_{j\in\mathcal{N}_i}\{\alpha_j^{-1}(\varepsilon_j) + \phi_{ij}\}$ ensures that (6.3.8) is satisfied as well, which concludes the proof. □

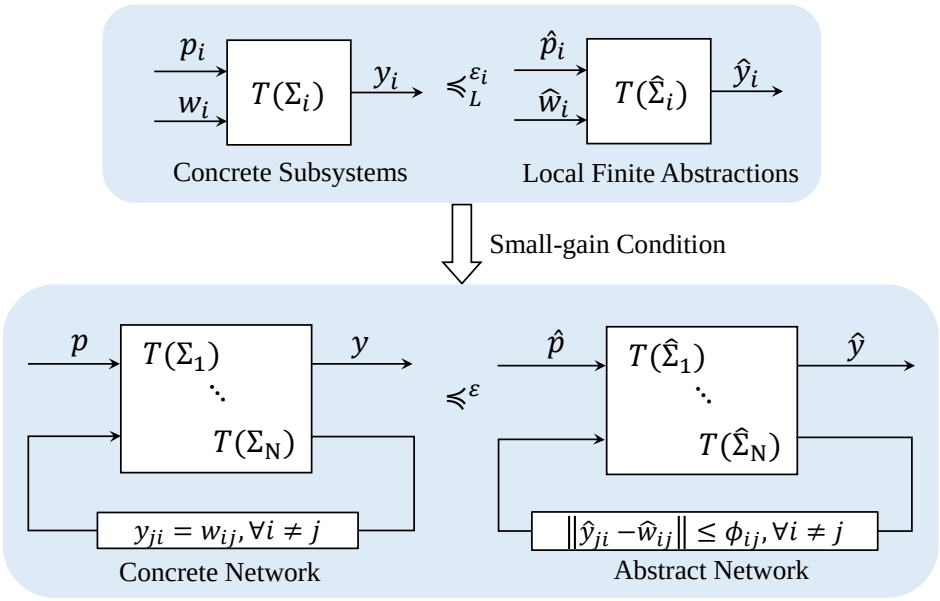


Figure 6.6: Compositional framework for the construction of opacity-preserving finite abstractions for interconnected switched systems.

Remark 6.3.29. *The compositionality result in Theorem 6.3.28 imposes a small-gain type condition on the concrete network of switched subsystems for the existence of proper modular finite abstraction, as depicted in Figure 6.6. In particular, under such small-gain type conditions, one can always find suitable local quantization parameters to construct local finite abstractions. The interconnection of the local finite abstractions can be used to serve as a finite abstraction for the concrete network satisfying the simulation relation $T(\Sigma) \preceq^{\varepsilon} T(\hat{\Sigma})$. Intuitively, the small-gain type condition facilitates the compositional construction of finite abstractions by certifying a small (weak) interaction of the subsystems which prevents an amplification of the signals across the possible interconnections.*

Remark 6.3.30. *Let us provide a general guideline on the computation of $\mathcal{K}_\infty$ functions σ_i, $i \in [1;N]$, that are used in Theorem 6.3.28: (i) in a general case when the network is consisting of $N \geq 1$ subsystems, functions $\sigma_i, i \in [1;N]$, can be constructed numerically by leveraging the algorithm introduced in [45] and the technique presented in [39, Proposition 8.8], see [158, Chapter 4]; (ii) for the case of having two and three subsystems in the network, there have been some construction techniques proposed in [77] and [39, Section 9], respectively; (iii) when the gain functions appeared in* (6.3.22) *satisfy $\gamma_{ij} < \mathrm{id}$, $\forall i,j \in [1;N]$, then one can always choose $\sigma_i, i \in [1;N]$ to be identity functions.*

Note that in this section we presented compositionality results on the construction of finite abstractions for notions of approximate initial-state and current-state opacity. One can readily follow the same lines of reasoning to establish similar results for the notion of approximate infinite-step opacity.

6.3.4 Case Study

Here, we provide an illustrative example to show how one can leverage the proposed compositional approach to check approximate initial-state opacity of a network of switched systems based on its finite abstraction.

Consider a network of discrete-time switched systems $\Sigma = (\mathbb{X}, \mathbb{X}_0, \mathbb{X}_S, P, F, Y, h)$ as in Definition 6.3.4, consisting of n subsystems Σ_i each described by:

$$\Sigma_i : \begin{cases} \mathbf{x}_i(k+1) &= a_{i\mathsf{p}_i(k)}\mathbf{x}_i(k) + d_i\omega_i(k) + b_{i\mathsf{p}_i(k)}, \\ \mathbf{y}_i(k) &= c_i\mathbf{x}_i(k), \end{cases} \tag{6.3.28}$$

where $\mathsf{p}_i(k) \in P_i = \{1,2\}$, $\forall k \in \mathbb{N}$, denotes the modes of each subsystem Σ_i. The other parameters are as the following: $a_{i1} = 0.05$, $a_{i2} = 0.1$, $b_{i1} = 0.1$, $b_{i2} = 0.15$, $d_i = 0.05$, $c_i = [c_{i1}; \ldots; c_{in}]$ with $c_{i(i+1)} = 1$, $c_{ij} = 0$, $\forall i \in [1;n-1], \forall j \neq i+1$, $c_{n1} = c_{nn} = 1$, $c_{nj} = 0$, $\forall j \in [2;n-1]$. The internal inputs are subject to the constraints $\omega_1(k) = c_{n1}\mathbf{x}_n(k)$ and $\omega_i(k) = c_{(i-1)i}\mathbf{x}_{(i-1)}(k)$, $\forall i \in [2;n]$. For each switched subsystem, the state set is $\mathbb{X}_i = \mathbb{X}_{0_i} = (0, 0.6)$, $\forall i \in [1;n]$, the secret set is $\mathbb{X}_{S_1} = (0, 0.2]$, $\mathbb{X}_{S_2} = [0.4, 0.6)$, $\mathbb{X}_{S_i} = (0, 0.6)$, $\forall i \in [3;n]$, the output set is $Y_i = \prod_{j=1}^{n} Y_{ij}$ where $Y_{i(i+1)} = (0, 0.6)$, $Y_{ii} = Y_{ij} = \{0\}$, $\forall i \in [1;n-1]$, $\forall j \neq i+1$, $Y_{nn} = Y_{n1} = (0, 0.6)$, $Y_{nj} = \{0\}$, $\forall j \in [2;n-1]$, and internal input set is $W_1 = Y_{ni}$, $W_i = Y_{(i-1)i}$, $\forall i \in [2;n]$. Intuitively, the output of the network is the external output of the last subsystem Σ_n. The interconnection topology of the network is depicted in Figure 6.7.

The main goal of this example is to check approximate initial-state opacity of the concrete network using its finite abstraction. Now, let us construct a finite abstraction of Σ compositionally with accuracy $\hat{\varepsilon} = 0.25$ as defined in (6.3.7), which preserves approximate initial-state opacity. We implement our compositional approach to achieve this goal.

Consider functions $V_{ip_i} = |x_i - \hat{x}_i|$, $\forall i \in [1;n]$. It can be readily verified that (6.3.10) and (6.3.11) are satisfied with $\underline{\alpha}_{ip_i} = \overline{\alpha}_{ip_i} = \mathrm{id}$, $\rho_{ip_i} = 0.05$, $\forall p_i \in P_i$, $\kappa_{i1} = a_{i1} = 0.05$, $\kappa_{i2} = a_{i2} = 0.1$. Condition (6.3.13) is satisfied with $\gamma_{ip_i} = \mathrm{id}$, $\forall p_i \in P_i$. Moreover, since

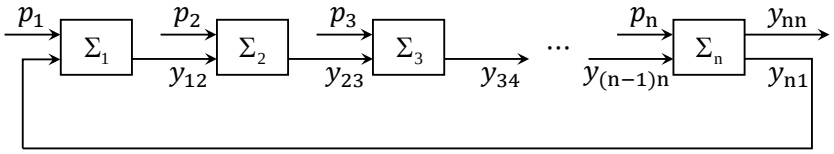


Figure 6.7: The interconnection topology of the network of discrete-time switched subsystems Σ_i.

$V_{ip_i} = V_{ip_i^+}, \forall p_i, p_i^+ \in P_i$, $V_i(x_i, \hat{x}_i) = |x_i - \hat{x}_i|$ is a common δ-ISS Lyapunov function for subsystem Σ_i. Next, given functions $\kappa_i = 0.1$, $\rho_i = 0.06\mathrm{id}$, $\alpha_i = \mathrm{id}$, $\hat{\gamma}_i = 1.05\mathrm{id}$, $\overline{\alpha}_i = \mathrm{id}$ as appeared in Theorem 6.3.24, we have $\gamma_{ij} < \mathrm{id}$ by (6.3.22), $\forall i, j \in [1; n]$. Hence, the small-gain condition (6.3.23) is satisfied. Then, by applying Theorem 6.3.28 and choosing functions $\sigma_i = \mathrm{id}$, $\forall i \in [1; n]$, such that (6.3.24) holds, we obtain proper pairs of local parameters $(\varepsilon_i, \vartheta_i) = (0.25, 0.25)$ for all of the transition systems. Accordingly, we provide a suitable choice of local quantization parameters as $\eta_i = 0.2$, $\forall i \in [1; n]$, such that inequality (6.3.14) for each transition system $T(\Sigma_i)$ is satisfied. Then, we construct local finite abstractions $T(\hat{\Sigma}_i) = (\hat{X}_i, \hat{X}_{0_i}, \hat{X}_{S_i}, \hat{U}_i, \hat{W}_i, \hat{\mathcal{F}}_i, \hat{Y}_i, \hat{\mathcal{H}}_i)$ as in Definition 6.3.21, where:

$$\hat{X}_i = \hat{X}_{0_i} = \{0.2, 0.4\}, \forall i \in [1; n],$$

$$\hat{X}_{S_i} = \begin{cases} \{0.2\}, & \text{if } i = 1 \\ \{0.4\}, & \text{if } i = 2 \\ \{0.2, 0.4\}, & \text{otherwise} \end{cases}$$

$$\hat{Y}_i = \begin{cases} \prod_{j=1}^{i} \{0\} \times \{0.2, 0.4\} \times \prod_{j=i+2}^{n} \{0\}, & \text{if } i \in [1; n-1] \\ \{0.2, 0.4\} \times \prod_{j=2}^{n-1} \{0\} \times \{0.2, 0.4\}, & \text{otherwise} \end{cases}$$

$$\hat{W}_i = \{0.2, 0.4\}, \forall i \in [1; n].$$

Using the result in Theorem 6.3.24, one can verify that $V_i(x_i, \hat{x}_i) = |x_i - \hat{x}_i|$ is a local ε_i-InitSOPSF from each $T(\Sigma_i)$ to its finite abstraction $T(\hat{\Sigma}_i)$. Furthermore, by the compositionality result in Theorem 6.3.17, we obtain that $V = \max_i \{V_i(x_i, \hat{x}_i)\} = \max_i \{|x_i - \hat{x}_i|\}$ is an ε-InitSOPSF from $T(\Sigma) = \mathcal{I}(T(\Sigma_1), \ldots, T(\Sigma_n))$ to $T(\hat{\Sigma}) = \hat{\mathcal{I}}(T(\hat{\Sigma}_1), \ldots, T(\hat{\Sigma}_n))$ with $\varepsilon = \max_i \varepsilon_i = 0.25$.

Now, let us verify approximate initial-state opacity for $T(\Sigma)$ using the network of finite abstractions $T(\hat{\Sigma})$. To do this, we first show an example of a network consisting of 3 transition systems, as shown in Figures 6.8 and 6.9. The three automata in Figure 6.8 represent the finite abstractions of the local transition systems, and the one in Figure 6.9 is the network of finite abstractions. Each circle is labeled by the state (top half) and the corresponding output (bottom half). Initial states are distinguished by being the target of a sourceless arrow. The states marked in red represent the secret states. The symbols on the edges show the switching signals $\mathsf{p}(k) \in \{1, 2\}^3$ and internal inputs coming from other local transition systems. For simplicity of demonstration, we use

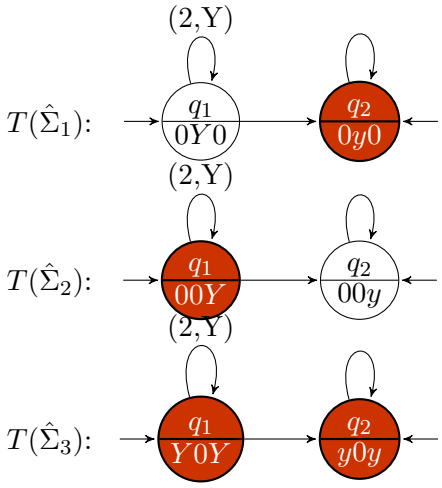


Figure 6.8: Local finite abstractions of transition systems.

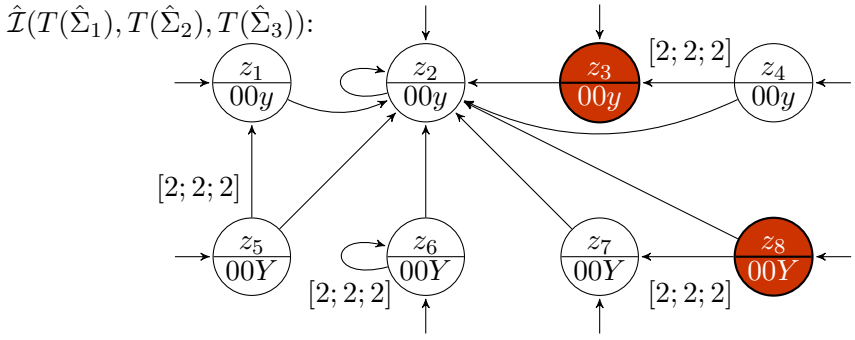


Figure 6.9: Finite abstraction of a network of 3 transition systems.

symbols to represent the state and output vectors, where the states of local transition systems are denoted by $q_1 = [0.4]$, $q_2 = [0.2]$, the states of network of transition systems are denoted by

$$z_1 = [q_1; q_2; q_2], z_2 = [q_2; q_2; q_2], z_3 = [q_2; q_1; q_2], z_4 = [q_1; q_1; q_2],$$
$$z_5 = [q_1; q_2; q_1], z_6 = [q_1; q_1; q_1], z_7 = [q_2; q_2; q_1], z_8 = [q_2; q_1; q_1],$$

and the outputs of the corresponding states are represented as $y = 0.2$ and $Y = 0.4$ with the symbols like $00y = [0; 0; 0.2]$, $00Y = [0; 0; 0.4]$ representing concatenated output vectors. One can easily see that $\hat{\mathcal{I}}(T(\hat{\Sigma}_1), T(\hat{\Sigma}_2), T(\hat{\Sigma}_3))$ is 0-approximate initial-state opaque, since for any run starting from any secret state, i.e. z_3 and z_8, there exists a run from a non-secret state, i.e. z_1 and z_6, such that the output trajectories are exactly the same. Essentially, one can verify that the abstract network holds this property regardless of the number of systems (i.e. n), due to the homogeneity of systems Σ_i and the symmetry of the circular network topology. Thus, one can conclude that $T(\hat{\Sigma}) = \hat{\mathcal{I}}(T(\hat{\Sigma}_1), \ldots, T(\hat{\Sigma}_n))$ is 0-approximate initial-state opaque. Therefore, by Corollary 6.3.14, we obtain that the original network $T(\Sigma) = \mathcal{I}(T(\Sigma_1), \ldots, T(\Sigma_n))$ is 0.5-approximate initial-state opaque.

6.4 A Barrier Certificate Approach for Interconnected Control Systems

As presented in Chapter 5, barrier certificates can be leveraged as a useful alternative approach for the verification of opacity for CPS. Though promising, the computational complexity of searching for parametric barrier certificates grow in polynomial time [198] with respect to the dimension of the system, and thus, the approaches proposed in Chapter 5 can become extremely expensive or even computationally intractable when dealing with large-scale interconnected control systems. In this section, we present a modular approach for verifying approximate opacity via the compositional construction of barrier certificates. This result shows that by employing a small-gain type condition, the desired augmented barrier certificates for an interconnected system can be constructed by composing so-called local barrier certificates of subsystems.

6.4.1 Augmented Control Subsystems

As discussed in Chapter 5, the verification of opacity for discrete-time control systems is achieved by computing barrier certificates defined over a so-called augmented system as in (5.2.1). Here, let us consider an interconnected system $\Sigma = (X, X_0, X_S, U, f, Y, h)$ composed of $N \in \mathbb{N}_{\geq 1}$ discrete-time control subsystems $\Sigma_i = (X_i, X_{0_i}, X_{S_i}, U_i, W_i, f_i, Y_i, h_i)$ as in Definition 6.2.2. We define the augmented system associated with a subsystem Σ_i by

$$\Sigma_i \times \Sigma_i = (X_i \times X_i, X_{0_i} \times X_{0_i}, X_{S_i} \times X_{S_i}, U_i \times U_i, W_i \times W_i, f_i \times f_i, Y_i \times Y_i, h_i \times h_i). \tag{6.4.1}$$

6.4.2 Compositional Construction of Barrier Certificates

In this subsection, we provide a compositional approach for the construction of barrier certificates to alleviate the computational cost encountered while dealing with large-scale interconnected systems. Here, we first show that by employing a small-gain type condition, a barrier certificate B for $\Sigma \times \Sigma$ as in Proposition 5.3.1 can be constructed by composing so-called local barrier certificates of subsystems as defined next.

Definition 6.4.1. *(Local barrier certificate for verifying opacity) Consider a control subsystem Σ_i. A function $B_i : X_i \times X_i \to \mathbb{R}$ is called a local barrier certificate for the augmented subsystem $\Sigma_i \times \Sigma_i$ if it satisfies the following conditions*

$$\forall (x_i, \hat{x}_i) \in \mathcal{R}_i, \qquad B_i(x_i, \hat{x}_i) \geq \alpha_i(||(h_i(x_i), h_i(\hat{x}_i))||), \tag{6.4.2}$$

$$\forall (x_i, \hat{x}_i) \in \mathcal{R}_{0_i}, \qquad B_i(x_i, \hat{x}_i) \leq \bar{\epsilon}_i, \tag{6.4.3}$$

$$\forall (x_i, \hat{x}_i) \in \mathcal{R}_{u_i}, \qquad B_i(x_i, \hat{x}_i) > \underline{\epsilon}_i, \tag{6.4.4}$$

$$\forall (x_i, \hat{x}_i) \in \mathcal{R}_i, \forall u_i \in U_i, \exists \hat{u}_i \in U_i, \forall (w_i, \hat{w}_i) \in W_i \times W_i,$$
$$B_i(f_i(x_i, u_i, w_i), f_i(\hat{x}_i, \hat{u}_i, \hat{w}_i)) \leq \kappa_i(B_i(x_i, \hat{x}_i)) + \gamma_{wi}(||(w_i, \hat{w}_i)||), \tag{6.4.5}$$

where sets $\mathcal{R}_{0_i}$ and $\mathcal{R}_{u_i}$ are the projections of sets $\mathcal{R}_0$ and $\mathcal{R}_u$ on the augmented subsystem $\Sigma_i \times \Sigma_i$, and α_i, γ_{wi}, $\kappa_i \in \mathcal{K}_\infty$, $\kappa_i \leq \mathrm{id}$, $\bar{\epsilon}_i, \underline{\epsilon}_i \in \mathbb{R}_{\geq 0}$.

Note that local barrier certificates of subsystems are mainly defined for constructing an overall barrier certificate for the interconnected system, and they are not useful on their own to verify opacity property. We now introduce the following lemma which will be used later in proving our main result.

Lemma 6.4.2. *For $a, b \in \mathbb{R}_{\geq 0}$, $\forall \lambda \in \mathcal{K}_\infty$, we have*

$$a + b \leq \max\{(\mathrm{id} + \lambda)(a), (\mathrm{id} + \lambda^{-1})(b)\}. \tag{6.4.6}$$

Proof. Define $c := \lambda^{-1}(b)$, we get the following

$$a + b = \begin{cases} a + \lambda(c) \leq c + \lambda(c) = (\mathrm{id} + \lambda^{-1})(b) & \text{if } a \leq c, \\ a + \lambda(c) < a + \lambda(a) = (\mathrm{id} + \lambda)(a) & \text{if } a > c, \end{cases}$$

which implies (6.4.6). □

For functions α_i, γ_{wi}, and κ_i associated with B_i as in Definition 6.4.1, we define, $\forall i, j \in [1; N]$,

$$\gamma_{ij} = \begin{cases} (\mathrm{id} + \lambda) \circ \kappa_i & \text{if } i = j, \\ (\mathrm{id} + \lambda^{-1}) \circ \gamma_{wi} \circ \alpha_j^{-1} & \text{if } i \neq j, \end{cases} \tag{6.4.7}$$

for some arbitrarily chosen $\lambda \in \mathcal{K}_\infty$.

Before stating our main compositionality result, we pose the following small-gain type assumption on the composition of gains γ_{ij}.

Assumption 6.4.3. *Assume functions γ_{ij} defined in* (6.4.7) *satisfy the following inequality*

$$\gamma_{i_1 i_2} \circ \gamma_{i_2 i_3} \circ \cdots \circ \gamma_{i_r i_1} < \mathrm{id}, \tag{6.4.8}$$

$\forall (i_1, \ldots, i_r) \in \{1, \ldots, N\}^r$, *where* $r \in \{1, \ldots, N\}$.

Note that by leveraging Theorem 5.2 in [39], the small gain condition in (6.4.8) *implies that there exists $\phi_i \in \mathcal{K}_\infty$, $\forall i \in [1; N]$, satisfying*

$$\max_{j \in [1;N]} \{\phi_i^{-1} \circ \gamma_{ij} \circ \phi_j\} < \mathrm{id}. \tag{6.4.9}$$

The following result shows that a barrier certificate B for the augmented interconnected system $\Sigma \times \Sigma$ can be obtained by composing local barrier certificates B_i computed for subsystems.

Theorem 6.4.4. *Consider an interconnected system $\Sigma = \mathcal{I}(\Sigma_1, \ldots, \Sigma_N)$, and the associated augmented system $\Sigma \times \Sigma$ composed of augmented subsystems $\Sigma_i \times \Sigma_i$. Assume each $\Sigma_i \times \Sigma_i$ admits a local barrier certificate B_i as in Definition 6.4.1. Let Assumption 6.4.3 hold, and $\max_{i \in [1;N]}\{\phi_i^{-1}(\overline{\epsilon}_i)\} \leq \max_{i \in [1;N]}\{\phi_i^{-1}(\underline{\epsilon}_i)\}$. Then, function $B : X \times X \to \mathbb{R}$ defined as*

$$B(x, \hat{x}) = \max_{i \in [1;N]} \{\phi_i^{-1} \circ B_i(x_i, \hat{x}_i)\}, \tag{6.4.10}$$

is a barrier certificate for $\Sigma \times \Sigma$ as in Proposition 5.3.1.

$$
\begin{aligned}
&B(f(x,u),f(\hat{x},\hat{u})) = \max_i\{\phi_i^{-1}\circ B_i(f_i(x_i,u_i,w_i),f_i(\hat{x}_i,\hat{u}_i,w_i))\}\\
&\overset{(6.4.5)}{\leq} \max_i\Big\{\phi_i^{-1}\big(\kappa_i(B_i(x_i,\hat{x}_i))+\gamma_{wi}(||(w_i,\hat{w}_i)||)\big)\Big\}\\
&\overset{(6.4.6)}{\leq} \max_i\Big\{\phi_i^{-1}\big(\max\{(\mathrm{id}+\lambda)(\kappa_i(B_i(x_i,\hat{x}_i))),(\mathrm{id}+\lambda^{-1})(\gamma_{wi}(||(w_i,\hat{w}_i)||))\}\big)\Big\}\\
&= \max_i\Big\{\phi_i^{-1}\big(\max\{(\mathrm{id}+\lambda)(\kappa_i(B_i(x_i,\hat{x}_i))),(\mathrm{id}+\lambda^{-1})(\gamma_{wi}(\max_{j,j\neq i}\{||(w_{ij},\hat{w}_{ij})||\}))\}\big)\Big\}\\
&= \max_i\Big\{\phi_i^{-1}\big(\max\{(\mathrm{id}+\lambda)(\kappa_i(B_i(x_i,\hat{x}_i))),(\mathrm{id}+\lambda^{-1})(\gamma_{wi}(\max_{j,j\neq i}\{||(y_{ji},\hat{y}_{ji})||\}))\}\big)\Big\}\\
&= \max_i\Big\{\phi_i^{-1}\big(\max\{(\mathrm{id}+\lambda)(\kappa_i(B_i(x_i,\hat{x}_i))),(\mathrm{id}+\lambda^{-1})(\gamma_{wi}(\max_{j,j\neq i}\{||(h_{ji}(x_j),h_{ji}(\hat{x}_j))||\}))\}\big)\Big\}\\
&\leq \max_i\Big\{\phi_i^{-1}\big(\max\{(\mathrm{id}+\lambda)(\kappa_i(B_i(x_i,\hat{x}_i))),(\mathrm{id}+\lambda^{-1})(\gamma_{wi}(\max_{j,j\neq i}\{||(h_j(x_j),h_j(\hat{x}_j))||\}))\}\big)\Big\}\\
&\overset{(6.4.2)}{\leq} \max_i\Big\{\phi_i^{-1}\big(\max\{(\mathrm{id}+\lambda)(\kappa_i(B_i(x_i,\hat{x}_i))),(\mathrm{id}+\lambda^{-1})(\gamma_{wi}(\max_{j,j\neq i}\{\alpha_j^{-1}\circ B_j(x_j,\hat{x}_j)\}))\}\big)\Big\}\\
&\overset{(6.4.7)}{\leq} \max_{i,j}\Big\{\phi_i^{-1}\circ\gamma_{ij}\circ B_j(x_j,\hat{x}_j)\Big\}\\
&\leq \max_{i,j,k}\Big\{\phi_i^{-1}\circ\gamma_{ij}\circ\phi_j\circ\phi_k^{-1}\circ B_k(x_k,\hat{x}_k)\Big\}\\
&\overset{(6.4.10)}{\leq} \max_{i,j}\Big\{\phi_i^{-1}\circ\gamma_{ij}\circ\phi_j\circ B(x,\hat{x})\Big\}\\
&\overset{(6.4.9)}{\leq} B(x,\hat{x}).
\end{aligned}
\tag{6.4.11}
$$

Proof. First, by Definition 6.4.1, we have

$$
\begin{aligned}
B(x,\hat{x}) &= \max_{i\in[1;N]}\{\phi_i^{-1}\circ B_i(x_i,\hat{x}_i)\} \overset{(6.4.3)}{\leq} \max_{i\in[1;N]}\{\phi_i^{-1}(\overline{\epsilon}_i)\},\\
B(x,\hat{x}) &= \max_{i\in[1;N]}\{\phi_i^{-1}\circ B_i(x_i,\hat{x}_i)\} \overset{(6.4.4)}{>} \max_{i\in[1;N]}\{\phi_i^{-1}(\underline{\epsilon}_i)\},
\end{aligned}
$$

which satisfies the first two conditions (5.3.1)-(5.3.2) in Proposition 5.3.1 by taking $\overline{\epsilon} = \max_{i\in[1;N]}\{\phi_i^{-1}(\overline{\epsilon}_i)\}$ and $\underline{\epsilon} = \max_{i\in[1;N]}\{\phi_i^{-1}(\underline{\epsilon}_i)\}$.

Next, by condition (6.4.5) of Definition 6.4.1, for all $(x,\hat{x})\in\mathcal{R}$ and $u\in U$, there exists $\hat{u}\in U$ such that $\forall(w_i,\hat{w}_i)\in W_i\times W_i$ the chain of inequalities in (6.4.11) holds. Recall that we set $w_{ij}=y_{ji}=h_{ji}(x_j)$ in (6.2.2) and (6.2.3). This gives us the identities in lines 4 and 5. The inequality in (6.4.11) satisfies the last condition (5.3.3) in Proposition 5.3.1. Therefore function B defined in (6.4.10) is a barrier certificate for the augmented interconnected system $\Sigma\times\Sigma$. □

Similarly, by applying the compositionality result proposed in Theorem 6.4.4, the barrier certificate V as in Proposition 5.3.2 for verifying the lack of opacity of an augmented system $\Sigma \times \Sigma$ can be computed by composing local barrier certificates V_i of subsystems as defined below.

Definition 6.4.5. *(Local barrier certificates for verifying lack of opacity) Consider a control subsystem Σ_i. A function $V_i : X_i \times X_i \to \mathbb{R}_{\geq 0}$ is called a local barrier certificate for the augmented subsystem $\Sigma_i \times \Sigma_i$ if it satisfies the following conditions*

$$\forall (x_i, \hat{x}_i) \in \mathcal{R}_i, \qquad V_i(x_i, \hat{x}_i) \geq \alpha_i(||(h_i(x_i), h_i(\hat{x}_i))||), \tag{6.4.12}$$

$$\forall (x_i, \hat{x}_i) \in \mathcal{R}_{0_i}, \qquad V_i(x_i, \hat{x}_i) \leq 0, \tag{6.4.13}$$

$$\forall (x_i, \hat{x}_i) \in \partial\mathcal{R}_i \setminus \partial\mathcal{R}_{u_i}, \qquad V_i(x_i, \hat{x}_i) > 0, \tag{6.4.14}$$

$$\begin{aligned} &\forall (x_i, \hat{x}_i) \in \overline{\mathcal{R}_i \setminus \mathcal{R}_{u_i}}, \forall u_i \in U_i, \exists \hat{u}_i \in U_i, \forall (w_i, \hat{w}_i) \in W_i \times W_i, \\ &\qquad V_i(f_i(x_i, w_i, u_i)) \leq \kappa_i(V_i(x_i, \hat{x}_i)) + \gamma_{wi}(||(w_i, \hat{w}_i)||), \end{aligned} \tag{6.4.15}$$

where the sets $\mathcal{R}_i$, $\mathcal{R}_{0_i}$, $\partial\mathcal{R}_i \setminus \partial\mathcal{R}_{u_i}$ and $\overline{\mathcal{R}_i \setminus \mathcal{R}_{u_i}}$ are, respectively, the projections of sets $\mathcal{R}$, $\mathcal{R}_0$, $\partial\mathcal{R} \setminus \partial\mathcal{R}_u$ and $\overline{\mathcal{R} \setminus \mathcal{R}_u}$ on the augmented subsystem $\Sigma_i \times \Sigma_i$, and α_i, γ_{wi}, $\kappa_i \in \mathcal{K}_\infty$, $\kappa_i \leq \mathrm{id}$, $\bar{\epsilon}_i, \underline{\epsilon}_i \in \mathbb{R}_{\geq 0}$.

Using the results of Theorem 6.4.4, one can construct a barrier certificate V for an augmented interconnected system $\Sigma \times \Sigma$, from the local barrier certificates V_i as in Definition 6.4.5.

6.4.3 Case Study

For systems with polynomial transition functions and semi-algebraic sets X_{0_i}, X_{S_i}, and X_i, we can use sum-of-squares (SOS) programming to search for polynomial local barrier certificates. We follow the same strategy as in Section 5.5, and use `SOSTOOLS` [138] together with a semidefinite programming solver `SeDuMi` [179] to compute local barrier certificates for subsystems in the following case study.

Consider a team of vehicles that are assigned to track a moving target. For the sake of simplicity, we constrain the target to move in a line, with bounded arbitrary acceleration. We also assume the vehicles are connected to each other in a line topology, and the distance between the first vehicle and the target is negligible. An intruder with δ measurement precision is trying to gain information about the initial position of the target. It has full knowledge of the system dynamics, but can only observe the position of the last vehicle in the team. Our aim is to verify whether the system is able to conceal its secret (defined as the initial position of the target) from the intruder. Figure 6.10 presents the experimental results of implementing our methodology for a team of $N = 100$ vehicles. The evolution of the states for the interconnected system is governed by

$$\begin{cases} \boldsymbol{\xi}_1(t+1) = A\boldsymbol{\xi}_1(t) + C\nu(t) + \boldsymbol{\xi}_2(t) \\ \boldsymbol{\xi}_2(t+1) = \nu(t) + \boldsymbol{\xi}_2(t) \end{cases} \tag{6.4.16}$$

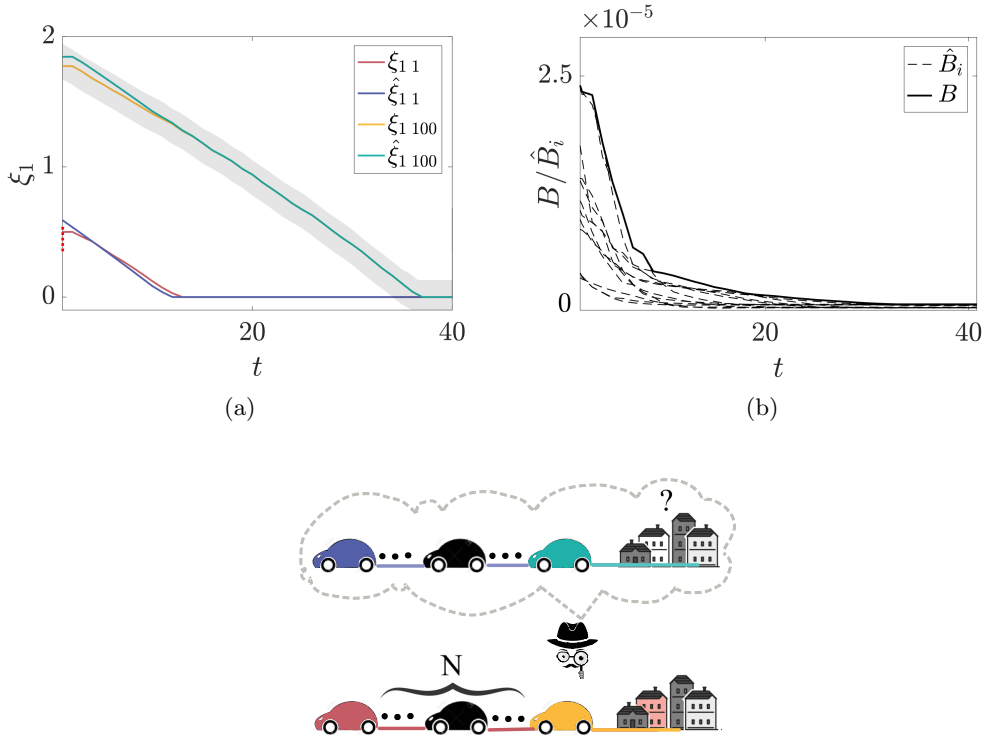


Figure 6.10: Results of simulating a system of 100 vehicles tracking a target. Target (red), and last vehicle (yellow) trajectories are plotted together with their corresponding non-secret pairs (blue and green lines). The shaded grey area indicates the region where the distance from the observed trajectory (yellow) is less than $\delta = 0.01$. The red dashed line on the x axis indicates the secret set for the target. b) The local barrier certificates computed for augmented subsystems (dashed lines), and their max in time, which is a barrier certificate for the interconnected system. c) The vehicles are connected together in a line topology, where vehicle i receives the position of $i-1$ as internal input. The intruder measures the location of the yellow vehicle, and tries to uncover the initial location of the target (red vehicle).

where $\boldsymbol{\xi}_1(t) = [\xi_{11}, \ldots, \xi_{1N}]^T \in \mathbb{R}^N$ and $\boldsymbol{\xi}_2(t) = [\xi_{21}, \ldots, \xi_{2N}]^T \in \mathbb{R}^N$ are the position and velocity vectors, respectively, and $\nu(t) \in \mathbb{R}^N$ contains the external input values of all the vehicles in the team. Taking $\mathbf{x}_i = [\xi_{1i}, \xi_{2i}]^T$, the following set of difference equations describe the dynamics of each subsystem Σ_i, $\forall i \in [1; N]$:

$$\Sigma_i : \begin{cases} \mathbf{x}_i(t+1) = \begin{bmatrix} 1-a & 1 \\ 0 & 1 \end{bmatrix} \mathbf{x}_i(t) + \nu_i(t) \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} + \mathbf{w}_i(t), \\ \mathbf{y}_i(t) = [0, \ldots, \mathbf{w}_{(i+1)i}(t), 0, \ldots, 0] \end{cases}$$

where $\mathbf{x}_0(t)$ is the state of the target at time t. In vector $\mathbf{w}_i(t)$, we have $\mathbf{w}_{i(i-1)}(t) = \mathbf{y}_{(i-1)i}(t) = \begin{bmatrix} a & 0 \end{bmatrix} \mathbf{x}_{i-1}(t)$, and all other entries are 0.

Matrix $A_{N\times N}$ in (6.4.16) represents the effects of internal input, as well as capturing the constant-acceleration motion of the vehicle i during the time interval $(t, t+1]$. Therefore, the entries A_{ij} are defined as $A_{ij} = \begin{cases} 1-a & \forall i = j, \\ a & \forall j = i-1, \end{cases}$ and zero else where. We set the constant $a = 0.01$ in this example. Matrix $C_{N\times N}$ is diagonal with $C_{i,i} = 0.5$, $\forall i \in [1; N]$. The output of the interconnected system is the position of the last vehicle, i.e., $\mathbf{y}(t) = [0, \ldots, 0, \xi_{1N}(t), 0]^T, N = 100$. The state set and initial set are $X = X_0 = \prod_{i=1}^{N} X_i$ where $X_i = X_{0_i} = [0, 2]$. The secret set for the interconnected system is set to $X_S = \prod_{i=1}^{N} X_{S_i}$, where $X_{S_1} = [0, 0.5]$, and X_{S_i} for all $i \in [2; 100]$ is a singleton containing a random number between $[0, 2]$. The measurement precision of the intruder is set to $\delta = 0.01$.

For finding local barrier certificates, we used $\bar{\epsilon}_i = 1, \underline{\epsilon}_i = 1.5$, for all $i \in [1; 100]$, $\alpha_j(r) = r$, $\kappa_i(r) = a\, r$, and $\gamma_{wi}(r) = a\, r, \forall r \in \mathbb{R}_{\geq 0}$. With the help of `SOSTOOLS` [138] and `SeDuMi` [179], we computed local barrier certificates together with their corresponding control policy $\hat{u}_i(\mathbf{x}_i, \hat{\mathbf{x}}_i, u_i) = \begin{bmatrix} 0.6 & -0.6 \end{bmatrix} \mathbf{x}_i + \begin{bmatrix} 1.2 & -1.2 \end{bmatrix} \hat{\mathbf{x}}_i + u_i$. One can readily verify that the small-gain assumption in (6.4.8) holds with $\gamma_{ij} < \mathrm{id}$, $\forall i, j \in [1; N]$. Then, by applying the results in Theorem 6.4.4, and taking $\phi_i = \mathrm{id}, \forall i \in [1; N]$, a barrier certificate for the interconnected system can be obtained as $B(x, \hat{x}) = \max_{i\in[1;N]}\{\hat{B}_i(x_i, \hat{x}_i)\}$. Figure 6.10(b) shows 10 of the computed local barrier certificates $\hat{B}_i$ for subsystems and the obtained overall barrier certificate B. The existence of the overall barrier certificate guarantees that for every state sequence of the interconnected system starting from a secret state, there exists at least another state sequence starting from a non-secret state such that the two sequences are indistinguishable in the eyes of the intruder with measurement precision δ. This is shown in Figure 6.10(a), where position sequences of the first and last vehicles (i.e. $\xi_{1\,1}$ and $\xi_{1\,100}$), are plotted with their corresponding $\hat{\xi}_{1\,1}$ and $\hat{\xi}_{1\,100}$, starting from non-secret initial states. One can readily see that the interconnected system is able to conceal its secret from possible intruders.

6.5 Discussion and Future Work

In this chapter, we proposed modular approaches for the verification of opacity for large-scale interconnected control (switched) systems. In Section 6.2, we proposed a methodology to compositionally construct opacity-preserving finite abstractions of interconnected discrete-time control systems. An approximate opacity-preserving simulation function is defined to characterize the simulation relation between two networks, which facilitates the abstraction-based opacity verification process. Then we presented a compositional approach to construct finite abstractions locally for concrete subsystems under incremental input-to-state stability property. The interconnection of local finite abstractions forms an abstract network that mimics the behaviors of the concrete network while preserving approximate opacity via the proposed simulation functions. Futhermore, we derived a small-gain type condition, under which one can guarantee

the existence of proper quantization parameters for the construction of finite abstractions. We further provided a top-down compositional construction framework along with a detailed quantization parameter design guideline. In Section 6.3, we further extended the compositional framework in Section 6.2 by enlarging the class of systems to hybrid ones with switching signals. An algorithm is also provided for the design of local quantization parameters tailored to networks of switched systems. In Section 6.4, we proposed an alternative modular approach via the notion of barrier certificates. Instead of leveraging abstraction-based techniques, we provided a compositional scheme for computing barrier certificates, which directly shows opacity of large-scale interconnected systems by finding local barrier certificates for their subsystems of much smaller and manageable sizes.

In the following, we further discuss some ongoing research topics and open problems on compositional approaches for opacity verification and synthesis.

Leverage Existing Modular Algorithms In the past decades, despite those opacity-related modular techniques already mentioned in Section 6.2, there are already numerous different modular verification and synthesis methods developed for other non-security properties in DES and formal methods literature. For example, in the context of supervisory control of DES, researchers have proposed many effective modular controller synthesis approaches using, for example, hierarchical interfaces [100, 68], state tree structures [123, 28], multi-level coordinators [92], and equivalence-based abstractions [46, 180]. There are also numerous recent works exploring the philosophy of compositional reasoning in the context of reactive synthesis; see, e.g., [5, 125, 14]. We believe that many of the aforementioned modular/compositional approaches for non-security properties can be generalized to incorporate the security constraints, which deserve deeper and detailed investigations.

Distributed Secure-by-Construction Synthesis For large-scale interconnected systems, the abstract interconnections constitute several relatively smaller local finite abstractions, as investigated in Section 6.2, that run synchronously. Since the controller synthesis problem for LTL specifications has severe worst-time complexity (doubly exponential), the product of all of the finite components makes the synthesis highly impractical. Moreover, often it may be impractical to assume that subsystems have complete knowledge of the states of other subsystems. To model these scenarios, one can represent the system as a network of finite abstractions where each subsystem has a separate mission and opacity requirement. Some of the states of neighbouring local finite abstractions may be shared with other local abstractions. This gives rise to the distributed reactive synthesis problem [170] where the system consists of several independent processes that cooperate based on local information to accomplish a global specification. Such a setting changes the synthesis problem from a two-player complete-information game to two-player games of incomplete information [156]. However, even for safety and reachability objectives (sub-classes of LTL), it is well known [141, 171] that the distributed synthesis problem is undecidable for general intercon-

nected systems. There are two directions to achieve decidability: the first is to restrict the network architecture [141] and the second is the approach of bounded synthesis [172].

7 Conclusions and Future Works

7.1 Conclusions

The main contribution of this dissertation is the development of theoretical foundations for the secure-by-construction design of cyber-physical systems. We conclude the dissertation by reviewing the results in the previous chapters.

In Chapter 3, we first reviewed some notions of security in three distinct research fields of discrete-event systems, control theory, and computer science. Despite the long history of security and privacy analysis in discrete-event systems and computer science communities, results on security analysis for cyber-physical systems are very limited. Based on these limitations, we developed a novel notion of security, called approximate opacity, which are more applicable to complex cyber-physical systems with possibly continuous state space. This new concept can be seen as a "robust" version of opacity by quantitatively characterizing the security guarantee level with respect to the measurement precision of the intruder. Moreover, in order to provide a general enough setting for the secure-by-construction synthesis scheme, we unified various security notions and mission requirements in a common framework, i.e., as a generalized language-based opacity notion. This allows us to develop cross-disciplinary theoretical results and leverage existing tools and algorithms from different research fields.

In Chapter 4, we focused on our proposed notion of approximate opacity and developed an abstraction-based framework for the verification of opacity for cyber-physical systems. In particular, we proposed new notions of approximate opacity-preserving simulation relations to capture the closeness between continuous-space systems and their finite abstractions (a.k.a symbolic models) in terms of preservation of approximate opacity. We also developed approaches for constructing opacity-preserving finite abstractions for a class of incrementally stable nonlinear control systems. Apart from this, we further extended the notion of approximate opacity to stochastic systems. Rather than providing a binary answer for opacity of stochastic systems (i.e., whether a system is opaque or not), a new notion of approximate opacity is introduced to quantitatively evaluate the possibility of a system being secure (i.e., the security level is quantified in a probabilistic setting). These works make the first step towards abstraction-based formal verification and synthesis of opacity. The developed abstraction-based scheme bridges the gap between opacity analysis of finite discrete systems and continuous (stochastic) control systems.

In Chapter 5, we developed a discretization-free framework as an alternative to the above abstraction-based method. For the first time, we proposed a deductive approach to formally verify opacity of continuous-space control systems using barrier certifi-

cates. Inspired by the duality of safety and reachability properties, a pair of so-called augmented control barrier certificates were defined for augmented systems that are constructed as the product of a control system and itself. The existence of the proposed barrier certificates are shown to guarantee the (or the lack of) opacity for general nonlinear control systems. Although both barrier certificates only serve as sufficient conditions, they can be utilized in reverse directions in the sense that one ensures approximate opacity, and the other one shows the lack of approximate opacity of control systems. We also presented a detailed way to compute polynomial barrier certificates using SOS programming under certain assumptions on the control systems.

In Chapter 6, modular approaches were proposed to alleviate the computational burden appeared in implementing the results proposed in the previous chapters. Note that although the abstraction-based approach provided in Chapter 4 and the deductive approach presented in Chapter 5 are shown to be promising tools, a challenge lies in scaling the approaches for large-scale systems. A particularly fruitful avenue to provide scalability is the compositional reasoning as discussed in Chapter 6. Here, a large-scale system is tackled as an interconnection of smaller subsystems with manageable sizes. We first presented a modular approach to reduce the computational complexity tailored to abstraction-based schemes. Instead of treating the interconnection monotonically, our compositionality result enables us to construct opacity-preserving finite abstractions for the subsystems individually. A top-down construction framework was presented with a detailed algorithm as a guideline for the design of quantization parameters. Here, we proposed compositionality results for both general interconnected control systems and interconnected switched systems which require different treatments on the state space discretization processes. Finally, a compositional scheme for the construction of barrier certificates is also derived based on a small-gain type condition. The proposed modular approaches are the first ones in the literature to analyze security notions of large-scale continuous-space systems.

7.2 Future Directions

Next, we touch upon some potential directions related to the secure-by-construction theme that differ from the parameters of study in this dissertation. We believe that these directions may provide impetus to research in security-critical system design.

Information-Theoretic Foundations The concept of privacy discussed so-far in this thesis is binary: either a system leaks information or it does not leak any information. However, in practice such binary mitigation may not be feasible and may require an information-theoretic prospective on quantifying and minimizing the amount of information leak. Shannon, in his seminal paper [175], coined and popularized the notion of *entropy* in measuring information: for a random variable X with values in some domain $\mathcal{X}$, the entropy of (or the uncertainty about) X, denoted by $H(X)$, is defined

as

$$H(X) = \sum_{x \in \mathcal{X}} P[X = x] \log_2 \frac{1}{P[X = x]}.$$

Shannon proved that $H(X)$ is the only function (modulo scaling) that satisfies the natural continuity, monotonicity, and choice decomposition (See [175], for more details). Similarly, for jointly distributed random variables X and Y, the conditional entropy $H(X \mid Y)$, i.e. uncertainty about X given Y, can be defined as

$$H(X \mid Y) = \sum_{y \in \mathcal{Y}} P[Y = y] H(X \mid Y = y),$$

where $\mathcal{Y}$ is the domain of Y. These definitions provide us a way to measure the *information loss*: if $H(X)$ is the uncertainty about X and if $H(X \mid Y)$ is the uncertainty about X after Y is revealed, the information loss in this process is $I(X;Y) = H(X) - H(X \mid Y)$. Smith [177] introduced an alternative notion of entropy called the *guessing entropy* $G(X)$ that corresponds to the number of guesses required to infer the value of X: of course a rational strategy in guessing these values will be to guess them in a non-increasing sequence of probability, hence $G(X) = \sum_{i=1}^{n} i p_i$ where $\langle p_1, p_2, \ldots, p_n \rangle$ is the sequence of probabilities of elements of X arranged in an non-increasing fashion.

The notion of opacity discussed in this thesis requires that the attacker should deduce nothing about all opacity properties of the system from observing the outputs of the system. However, achieving full opacity may not be possible in general, because oftentimes systems reveal information depending on the secret properties. To extend the notion of opacity to quantitative opacity, we can use the quantitative notion of information leakage. We say that two opacity properties α, α' are *indistinguishable* in Σ, and we write $\alpha \equiv_\Sigma \alpha'$, if for any trace r satisfying α, there exists another trace r' satisfying α' such that both r and r' have analogous observations, i.e. $h(r) = h(r')$. Let us generalize the original set of opacity properties from $\{\alpha, \neg\alpha\}$ to $\overline{\alpha} = \{\alpha_1, \ldots, \alpha_n\}$. In this case, the system Σ is called *opaque*, if every pair of opacity properties in $\overline{\alpha}$ are mutually indistinguishable. Let $Q = \{Q_1, Q_2, \ldots, Q_k\}$ be the quotient space of O characterized by the indistinguishability relation. Let $B_Q = \langle B_1, B_2, \ldots, B_k \rangle$ be the sizes of observational equivalence classes from Q; let $B = \sum_{i=1}^{k} B_i$. Assuming uniform distributions on Q, Köpf and Basin [93] characterize expressions for various information-theoretic measures on information leaks which are given below:

1. *Shannon Entropy:* $SE(\Sigma, \overline{\alpha}) = (\frac{1}{B}) \sum\limits_{1 \le i \le k} B_i \log_2(B_i)$,
2. *Guessing Entropy:* $GE(\Sigma, \overline{\alpha}) = (\frac{1}{2B}) \sum\limits_{1 \le i \le k} B_i^2 + \frac{1}{2}$,
3. *Min-Guess Entropy:* $MG(\Sigma, \overline{\alpha}) = \min\limits_{1 \le i \le k} \{(B_i + 1)/2\}$.

This allows us to generalize our opacity requirements in a quantitative fashion. Given a property φ as a mission requirement, and opacity property tuple $\overline{\alpha} = \{\alpha_1, \ldots, \alpha_k\}$, an entropy bound K and the corresponding entropy criterion $\kappa \in \{SE, GE, MG\}$,

the quantitative security-aware verification $\Sigma \models (\varphi, \overline{\alpha})$ is to decide whether $\Sigma \models \varphi$ and $\kappa(\Sigma, \overline{\alpha}) \leq K$. Similarly, the quantitative security-aware synthesis is to design a supervisor/controller C such that $\Sigma_C \models (\varphi, \overline{\alpha})$.

Quantitative theory of information have been widely used for the verification of security properties [177, 93, 12, 67] in the context of finite state and software systems. Moreover, for such systems several restricted classes of synthesis approaches [94, 11, 220, 221, 80, 174, 189] have been proposed that focus on side-channel mitigation techniques by increasing the remaining entropy of secret sets leaked while maintaining the performance.

Data-Driven Approaches for CPS Security This thesis assumed the access to a model of the system and proposed security-aware verification and synthesis approaches. Oftentimes, a true explicit model of the system is not available or is too large to reason with formally. Reinforcement learning [181] (RL) is a sampling-based optimization algorithm that computes optimal policies driven by scalar reward signals. Recently, RL has been extended to work with formal logic [23, 24, 137, 64, 96], and automatic structures (ω-automata [60, 61] and reward machines [74]) instead of scalar reward signals. A promising future direction is to extend RL-based synthesis to reason with security properties of the system.

The controller learned via deep RL will have deep neural networks as the controllers. Additionally, deep neural networks are often employed in place of cumbersome tabular controllers to minimize the size of the program logic. In such systems, security verification need to reason with neural networks along with the system dynamics. There is a large body of work [71, 1, 60, 151, 119, 203, 96] in verifying control systems with neural networks using SMT solvers, and will provide a promising avenue of research in developing security verification and synthesis approaches for CPS with neural networks based controllers.

Radical advances in inexpensive sensors, wireless technology, and the Internet of Things (IoT) offer unprecedented opportunities by ubiquitously collecting data at high detail and at large scale. Utilization of data at these scales, however, poses a major challenge for verifying or designing CPS, particularly in view of the additional inherent uncertainty that data-driven signals introduce to systems behavior and their correctness. In fact, this effect has not been rigorously understood to this date, primarily due to the missing link between data analytics techniques in machine learning/optimization and the underlying physics of CPS. A future research direction is to develop scalable data-driven approaches for formal verification and synthesis of CPS with unknown closed form models (a.k.a. black-box systems) with respect to both mission and security properties. The main novelty is to bypass the model identification phase and directly verify or synthesize controller for CPS using system behaviors. The main reasons behind the quest to directly work on system behaviors and bypass the identification phase are: i) Identification can introduce approximation errors and have a large computational complexity; ii) Even when the model is known, formal verification and synthesis of CPS are computationally challenging.

Security for Network Multi-Agent CPS This thesis mostly discussed a centralized setting for CPS security, i.e., a single CPS plant with global secrets against a single attacker, although the CPS itself may consist of several smaller subsystems. However, in many modern engineering systems such as connected autonomous vehicles [120], smart micro-grids [213] and smart cities [25], there may exist no centralized decision-maker. Instead, each CPS agent interacts and collaborates/competes with each other via information exchanges over networks to make decisions, which leads to the network multi-agent CPS. There is a large body of works [55, 194, 56, 83, 173, 166] in synthesizing coordination strategies for network multi-agent CPS for high-level mission requirements using formal methods. However, the security issue, which is more severe in multi-agent CPS due to large communications and information exchanges, is rarely considered. In particular, in multi-agent CPS, each agent may have its own security considerations that depend on the time-varying configurations of the entire network. Therefore, how to define formal security notions that are suitable for multi-agent systems is an important but challenging future direction.

Recently, security and privacy considerations over networks have attracted significant attentions in the context of distributed state estimations [129, 8], distributed averaging/consensus [130, 58], distributed optimizations [63, 121], and distributed machine learning [72, 105]. However, those results are mostly developed for distributed computing systems and are not directly applicable for multi-agent CPS with heterogeneous dynamics. Furthermore, most of the existing security-aware protocols for distributed systems are designed for specific tasks and there is still a lack of formal methodologies for security-aware verification and secure-by-construction synthesis of communication protocols and coordination strategies for network multi-agent CPS. Finally, rather than a single passive attacker, network CPS may suffer from multiple active malicious attackers. Therefore, one needs to develop effective approaches for characterizing and controlling the evolution of security properties over dynamic networks of multiple players. A promising future direction is to develop a comprehensive framework for multi-agent CPS security by extending formal reasoning with multi-player game-theory.

We shall draw the readers' and potential researchers' attention that, security has been a moving goalpost and more damaging vulnerabilities are yet unknown. The proposed approaches in this dissertation need to be combined with classical fuzzing-based security research to uncover previously undiscovered security vulnerabilities. Moreover, most of the existing results on security analysis for CPS remain mainly theoretical. There is a great need to develop efficient toolboxes and proof-of-concept benchmarks to evaluate the practical feasibility of the foundations and algorithms developed for abstracting, analyzing, or enforcing security properties over complex CPS. In addition to academic benchmarks, it is important to improve the applicability of theoretical methods to industrial case studies and real-life applications.

Bibliography

[1] A. Abate, D. Ahmed, M. Giacobbe, and A. Peruffo. Formal synthesis of Lyapunov neural networks. *IEEE Control Systems Letters*, 5(3):773–778, 2021.

[2] M. Ahmadi, B. Wu, H. Lin, and U. Topcu. Privacy verification in POMDPs via barrier certificates. In *57th IEEE Conference on Decision and Control (CDC)*, pages 5610–5615, 2018.

[3] R. Alur, P. Černý, and S. Zdancewic. Preserving secrecy under refinement. In *Automata, Languages and Programming*, pages 107–118. Springer Berlin Heidelberg, 2006.

[4] R. Alur, T. Henzinger, G. Lafferriere, and G. J. Pappas. Discrete abstractions of hybrid systems. *Proceedings of the IEEE*, 88(7):971–984, 2000.

[5] R. Alur, S. Moarref, and U. Topcu. Compositional and symbolic synthesis of reactive controllers for multi-agent systems. *Information and Computation*, 261:616–633, 2018.

[6] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada. Control barrier functions: Theory and applications. In *18th European Control Conference (ECC)*, pages 3420–3431, 2019.

[7] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada. Control barrier function based quadratic programs for safety critical systems. *IEEE Transactions on Automatic Control*, 62(8):3861–3876, 2017.

[8] L. An and G.-H. Yang. Enhancement of opacity for distributed state estimation in cyber–physical systems. *Automatica*, 136:110087, 2022.

[9] M. Anand, V. Murali, A. Trivedi, and M. Zamani. Formal verification of control systems against hyperproperties via barrier certificates. *arXiv preprint arXiv:2105.05493*, 2021.

[10] D. Angeli. A Lyapunov approach to incremental stability properties. *IEEE Transactions on Automatic Control*, 47(3):410–21, 2002.

[11] A. Askarov, D. Zhang, and A. C. Myers. Predictive black-box mitigation of timing channels. In *Proceedings of the 17th ACM conference on Computer and communications security*, pages 297–307, 2010.

[12] M. Backes, B. Köpf, and A. Rybalchenko. Automatic discovery and quantification of information leaks. In *30th IEEE Symposium on Security and Privacy*, pages 141–153, 2009.

[13] C. Baier and J. P. Katoen. *Principles of model checking.* The MIT Press, 2008.

[14] G. Bakirtzis, E. Subrahmanian, and C. H. Fleming. Compositional thinking in cyberphysical systems theory. *Computer*, 54(12):50–59, 2021.

[15] G. Barequet and S. Har-Peled. Efficiently approximating the minimum-volume bounding box of a point set in three dimensions. *Journal of Algorithms*, 38(1):91–109, 2001.

[16] F. Basile and G. De Tommasi. An algebraic characterization of language-based opacity in labeled Petri nets. In *14th International Workshop on Discrete Event Systems*, pages 329–336, 2018.

[17] C. Belta, B. Yordanov, and E. Göl. *Formal Methods for Discrete-Time Dynamical Systems*, volume 89. Springer International Publishing, 2017.

[18] B. Bérard, K. Chatterjee, and N. Sznajder. Probabilistic opacity for Markov decision processes. *Information Processing Letters*, 115(1):52–59, 2015.

[19] B. Bérard, J. Mullins, and M. Sassolas. Quantifying opacity. *Mathematical Structures in Computer Science*, 25(2):361–403, 2015.

[20] D. Bestvater, E. V. Dunn, C. Townsend, and W. Nelson. Satisfaction and wait time of patients visiting a family practice clinic. *Canadian family physician (Medecin de famille canadien)*, 34:67–70, 1988.

[21] A. Borri, G. Pola, and M. D. Di Benedetto. Design of symbolic controllers for networked control systems. *IEEE Transactions on Automatic Control*, 64(3):1034–1046, 2019.

[22] D. Boskos and D. V. Dimarogonas. Decentralized abstractions for feedback interconnected multi-agent systems. In *54th IEEE Conference on Decision and Control (CDC)*, pages 282–287, 2015.

[23] A. Camacho, O. Chen, S. Sanner, and S. A. McIlraith. Non-Markovian rewards expressed in LTL: guiding search via reward shaping. In *Tenth Annual Symposium on Combinatorial Search*, 2017.

[24] A. Camacho, R. T. Icarte, T. Q. Klassen, R. A. Valenzano, and S. A. McIlraith. LTL and beyond: Formal languages for reward function specification in reinforcement learning. In *International Joint Conferences on Artificial Intelligence Organization (IJCAI)*, volume 19, pages 6065–6073, 2019.

[25] C. G. Cassandras. Smart cities as cyber-physical social systems. *Engineering*, 2(2):156–158, 2016.

[26] C. G. Cassandras and S. Lafortune. *Introduction to discrete event systems*, volume 3. Springer, 2021.

[27] 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.

[28] W. Chao, Y. Gan, Z. Wang, and W. M. Wonham. Modular supervisory control and coordination of state tree structures. *International Journal of Control*, 86(1):9–21, 2013.

[29] 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*, 25(1-2):271–294, 2014.

[30] J. Chen, M. Ibrahim, and R. Kumar. Quantification of secrecy in partially observed stochastic discrete event systems. *IEEE Trans. Automation Science and Engineering*, 14(1):185–195, 2017.

[31] D. Chistikov, A. S. Murawski, and D. Purser. Asymmetric distances for approximate differential privacy. In *30th International Conference on Concurrency Theory (CONCUR 2019)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

[32] M. R. Clarkson, B. Finkbeiner, M. Koleini, K. K. Micinski, M. N. Rabe, and C. Sánchez. Temporal logics for hyperproperties. In *Principles of Security and Trust*, pages 265–284, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.

[33] M. R. Clarkson and F. B. Schneider. Hyperproperties. *Journal of Computer Security*, 18(6):1157–1210, 2010.

[34] L. Clavijo and J. Basilio. Empirical studies in the size of diagnosers and verifiers for diagnosability analysis. *Discrete Event Dynamic Systems*, 27(4):701–739, 2017.

[35] X. Cong, M. P. Fanti, A. M. Mangini, and Z. Li. On-line verification of current-state opacity by petri nets and integer linear programming. *Automatica*, 94:205–213, 2018.

[36] C. Cortes, M. Mohri, and A. Rastogi. Lp distance and equivalence of probabilistic automata. *International Journal of Foundations of Computer Science*, 18(04):761–779, 2007.

[37] M. Coste. An introduction to semialgebraic geometry, 2000.

[38] S. Dashkovskiy, B. S. Rüffer, and F. R. Wirth. An ISS small gain theorem for general networks. *Mathematics of Control, Signals, and Systems*, 19(2):93–122, 2007.

[39] S. N. Dashkovskiy, B. S. Rüffer, and F. R. Wirth. Small gain theorems for large scale systems and construction of ISS Lyapunov functions. *SIAM Journal on Control and Optimization*, 48(6):4089–4118, 2010.

[40] G. De Giacomo and M. Y. Vardi. Linear temporal logic and linear dynamic logic on finite traces. In *23rd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 854–860. AAAI Press, 2013.

[41] L. De Moura and N. Bjørner. Z3: An efficient SMT solver. In *International conference on Tools and Algorithms for the Construction and Analysis of Systems*, pages 337–340. Springer, 2008.

[42] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labelled markov processes. *Theoretical computer science*, 318(3):323–354, 2004.

[43] J. Ding, M. Kamgarpour, S. Summers, A. Abate, J. Lygeros, and C. Tomlin. A stochastic games framework for verification and control of discrete time stochastic hybrid systems. *Automatica*, 49(9):2665–2674, 2013.

[44] C. Dwork. Differential privacy. *Encyclopedia of Cryptography and Security*, pages 338–340, 2011.

[45] B. C. Eaves. Homotopies for computation of fixed points. *Mathematical Programming*, 3(1):1–22, 1972.

[46] L. Feng and W. M. Wonham. Supervisory control architecture for discrete-event systems. *IEEE Transactions on Automatic Control*, 53(6):1449–1461, 2008.

[47] G. Fiore, E. De Santis, G. Pola, and M. Di Benedetto. On approximate predictability of metric systems. In *6th IFAC Conference on Analysis and Design of Hybrid Systems*, pages 169–174, 2018.

[48] S. Gao, S. Kong, and E. M. Clarke. dReal: An SMT solver for nonlinear theories over the reals. In *International conference on automated deduction*, pages 208–214. Springer, 2013.

[49] D. Genkin, A. Shamir, and E. Tromer. Rsa key extraction via low-bandwidth acoustic cryptanalysis. In *Advances in Cryptology – CRYPTO*, pages 444–461. Springer Berlin Heidelberg, 2014.

[50] A. Girard, A. A. Julius, and G. J. Pappas. Approximate simulation relations for hybrid systems. *Discrete event dynamic systems*, 18(2):163–179, 2008.

[51] A. Girard and G. J. Pappas. Approximation metrics for discrete and continuous systems. *IEEE Transactions on Automatic Control*, 52(5):782–798, 2007.

[52] A. Girard and G. J. Pappas. Approximation metrics for discrete and continuous systems. *IEEE Transactions on Automatic Control*, 25(5):782–798, 2007.

[53] A. Girard, G. Pola, and P. Tabuada. Approximately bisimilar symbolic models for incrementally stable switched systems. *IEEE Transactions on Automatic Control*, 55(1):116–126, 2010.

[54] A. Greenberg. Hackers remotely kill a jeep on the highway—with me in in. `https://www.wired.com/2015/07/hackers-remotely-kill-jeep-highway/`, 2015. Online published 21-July-2015.

[55] M. Guo and D. V. Dimarogonas. Multi-agent plan reconfiguration under local LTL specifications. *The International Journal of Robotics Research*, 34(2):218–235, 2015.

[56] M. Guo, J. Tumova, and D. V. Dimarogonas. Communication-free multi-agent control under local temporal tasks and relative-distance constraints. *IEEE Transactions on Automatic Control*, 61(12):3948–3962, 2016.

[57] C. N. Hadjicostis. *Estimation and Inference in Discrete Event Systems*. Springer, 2020.

[58] C. N. Hadjicostis and A. D. Domínguez-García. Privacy-preserving distributed averaging via homomorphically encrypted ratio consensus. *IEEE Transactions on Automatic Control*, 65(9):3887–3894, 2020.

[59] S. Haesaert, S. Soudjani, and A. Abate. Verification of general markov decision processes by approximate similarity relations and policy refinement. *SIAM Journal on Control and Optimization*, 55(4):2333–2367, 2017.

[60] E. M. Hahn, M. Perez, S. Schewe, F. Somenzi, A. Trivedi, and D. Wojtczak. Omega-regular objectives in model-free reinforcement learning. In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, pages 395–412. Springer, 2019.

[61] E. M. Hahn, M. Perez, S. Schewe, F. Somenzi, A. Trivedi, and D. Wojtczak. Model-free reinforcement learning for lexicographic ω-regular objectives. *International Symposium on Formal Methods*, 2021.

[62] D. Halperin, T. S. Heydt-Benjamin, B. Ransford, S. S. Clark, B. Defend, W. Morgan, K. Fu, T. Kohno, and W. H. Maisel. Pacemakers and implantable cardiac defibrillators: Software radio attacks and zero-power defenses. In *IEEE Symposium on Security and Privacy*, pages 129–142, 2008.

[63] S. Han, U. Topcu, and G. J. Pappas. Differentially private distributed constrained optimization. *IEEE Transactions on Automatic Control*, 62(1):50–64, 2017.

[64] M. Hasanbeig, A. Abate, and D. Kroening. Certified reinforcement learning with logic guidance. *arXiv preprint arXiv:1902.00778*, 2019.

[65] K. Hashimoto, A. Saoud, M. Kishida, T. Ushio, and D. V. Dimarogonas. A symbolic approach to the self-triggered design for networked control systems. *IEEE Control Systems Letters*, 3(4):1050–1055, 2019.

[66] R. Hermann and A. Krener. Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, 22(5):728–740, 1977.

[67] J. Heusser and P. Malacaria. Quantifying information leaks in software. In *Proceedings of the 26th Annual Computer Security Applications Conference*, pages 261–269. ACM, 2010.

[68] R. C. Hill, J. E. R. Cury, M. H. de Queiroz, D. M. Tilbury, and S. Lafortune. Multi-level hierarchical interface-based supervisory control. *Automatica*, 46(7):1152–1164, 2010.

[69] J. Hou, S. Liu, X. Yin, and M. Zamani. Abstraction-based verification of approximate pre-opacity for control systems. *arXiv preprint arXiv:2211.04098*, 2022.

[70] J. Hou, X. Yin, S. Li, and M. Zamani. Abstraction-based synthesis of opacity-enforcing controllers using alternating simulation relations. In *58th IEEE Conference on Decision and Control (CDC)*, pages 7653–7658, 2019.

[71] X. Huang, M. Kwiatkowska, S. Wang, and M. Wu. Safety verification of deep neural networks. In *International conference on computer aided verification (CAV)*, pages 3–29. Springer, 2017.

[72] Y. Huang, Z. Song, K. Li, and S. Arora. Instahide: Instance-hiding schemes for private distributed learning. In *International Conference on Machine Learning*, pages 4507–4518, 2020.

[73] M. Hutter and J.-M. Schmidt. The temperature side-channel and heating fault attacks. In *International Conference on Smart Card Research and Advanced Applications*, volume 8419, pages 219–235. Springer, 2013.

[74] R. T. Icarte, T. Klassen, R. Valenzano, and S. McIlraith. Using reward machines for high-level task specification and decomposition in reinforcement learning. In *International Conference on Machine Learning*, pages 2107–2116, 2018.

[75] R. Jacob, J.-J. Lesage, and J.-M. Faure. Overview of discrete event systems opacity: Models, validation, and quantification. *Annual Reviews in Control*, 41:135–146, 2016.

[76] P. Jagtap, S. Soudjani, and M. Zamani. Formal synthesis of stochastic systems via control barrier certificates. *IEEE Transactions on Automatic Control*, 66(7):3097–3110, 2020.

[77] Z. Jiang, I. M. Y. Mareels, and Y. Wang. A Lyapunov formulation of the nonlinear small-gain theorem for interconnected ISS systems. *Automatica*, 32(8), 1996.

[78] Z.-P. Jiang, A. R. Teel, and L. Praly. Small-gain theorem for ISS systems and applications. *Mathematics of Control, Signals and Systems*, 7(2):95–120, 1994.

[79] A. Jones, K. Leahy, and M. Hale. Towards differential privacy for symbolic systems. *arXiv preprint arXiv:1809.08634*, 2018.

[80] S. Kadloor, N. Kiyavash, and P. Venkitasubramaniam. Mitigating timing based information leakage in shared schedulers. In *Proceedings IEEE INFOCOM*, pages 1044–1052, 2012.

[81] S. T. Kalat, S. Liu, and M. Zamani. Modular verification of opacity for interconnected control systems via barrier certificates. *IEEE Control Systems Letters*, 6:890–895, 2021.

[82] S. T. Kalat, S. Liu, and M. Zamani. Verification of approximate infinite-step opacity using barrier certificates. In *European Control Conference (ECC)*, pages 175–180, 2022.

[83] Y. Kantaros and M. M. Zavlanos. Distributed intermittent connectivity control of mobile robot networks. *IEEE Transactions on Automatic Control*, 62(7):3109–3121, 2016.

[84] C. Keroglou and C. N. Hadjicostis. Probabilistic system opacity in discrete event systems. *Discrete Event Dynamic Systems*, 28(2):289–314, 2018.

[85] M. Khaled, K. Zhang, and M. Zamani. Output-feedback symbolic control. *arXiv preprint arXiv:2011.14848*, 2020.

[86] S. Kiefer. On computing the total variation distance of hidden markov models. *arXiv preprint arXiv:1804.06170*, 2018.

[87] E. S. Kim, M. Arcak, and S. A. Seshia. Compositional controller synthesis for vehicular traffic networks. In *54th IEEE Conference on Decision and Control (CDC)*, pages 6165–6171, 2015.

[88] E. S. Kim, M. Arcak, and S. A. Seshia. A small gain theorem for parametric assume-guarantee contracts. In *International Conference on Hybrid Systems: Computation and Control (HSCC)*, pages 207–216. ACM, 2017.

[89] E. S. Kim, M. Arcak, and M. Zamani. Constructing control system abstractions from modular components. In *21st International Conference on Hybrid Systems: Computation and Control (HSCC)*, pages 137–146. ACM, 2018.

[90] K.-D. Kim and P. Kumar. Cyber–physical systems: A perspective at the centennial. *Proceedings of the IEEE*, 100(Special Centennial Issue):1287–1308, 2012.

[91] K. Kobayashi and K. Hiraishi. Verification of opacity and diagnosability for pushdown systems. *J. Applied Mathematics*, 2013, 2013.

[92] J. Komenda, T. Masopust, and J. H. van Schuppen. Coordination control of discrete-event systems revisited. *Discrete Event Dynamic Systems*, 25(1):65–94, 2015.

[93] B. Köpf and D. Basin. An information-theoretic model for adaptive side-channel attacks. In *14th ACM Conference on Computer and Communications Security*, pages 286–296, New York, NY, USA, 2007.

[94] B. Köpf and M. Dürmuth. A provably secure and efficient countermeasure against timing attacks. In *22nd IEEE Symposium on Computer Security Foundations*, pages 324–335, 2009.

[95] S. Lafortune, F. Lin, and C. N. Hadjicostis. On the history of diagnosability and opacity in discrete event systems. *Annual Reviews in Control*, 45:257–266, 2018.

[96] A. Lavaei, F. Somenzi, S. Soudjani, A. Trivedi, and M. Zamani. Formal controller synthesis for continuous-space MDPs via model-free reinforcement learning. In *11th International Conference on Cyber-Physical Systems (ICCPS)*, pages 98–107. IEEE, 2020.

[97] A. Lavaei, S. Soudjani, and M. Zamani. From dissipativity theory to compositional construction of finite Markov decision processes. In *International Conference on Hybrid Systems: Computation and Control (HSCC)*, pages 21–30. ACM, 2018.

[98] A. Lavaei, S. Soudjani, and M. Zamani. Compositional (in) finite abstractions for large-scale interconnected stochastic systems. *IEEE Transactions on Automatic Control*, 65(12):5280–5295, 2020.

[99] J. Le Ny and G. Pappas. Differentially private filtering. *IEEE Transactions on Automatic Control*, 59(2):341–354, 2014.

[100] R. J. Leduc, B. A. Brandin, M. Lawford, and W. M. Wonham. Hierarchical interface-based supervisory control-part i: serial case. *IEEE Transactions on Automatic Control*, 50(9):1322–1335, 2005.

[101] E. A. Lee and S. A. Seshia. *Introduction to embedded systems, a cyber-physical systems approach*. MIT Press, second edition, 2017.

[102] J.-W. Lee and G. E. Dullerud. Uniform stabilization of discrete-time switched and markovian jump linear systems. *Automatica*, 42(2):205–218, 2006.

[103] D. Lefebvre and C. N. Hadjicostis. Exposure and revelation times as a measure of opacity in timed stochastic discrete event systems. *IEEE Transactions on Automatic Control*, 66(12):5802–5815, 2020.

[104] P. Leu, I. Puddu, A. Ranganathan, and S. Čapkun. I send, therefore I leak: Information leakage in low-power wide area networks. In *11th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, pages 23–33, 2018.

[105] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith. Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3):50–60, 2020.

[106] D. Liberzon. *Switching in Systems and Control*. Birkhäuser Basel, 2003.

[107] F. Lin. Opacity of discrete event systems and its applications. *Automatica*, 47(3):496–503, 2011.

[108] F. Lin, L. Y. Wang, W. Chen, W. Wang, and F. Wang. Information control in networked discrete event systems and its application to battery management systems. *Discrete Event Dynamic Systems*, 30(2):243–268, 2020.

[109] L. Lindemann and D. V. Dimarogonas. Control barrier functions for signal temporal logic tasks. *IEEE control systems letters*, 3(1):96–101, 2018.

[110] S. Liu, N. Noroozi, and M. Zamani. Symbolic models for infinite networks of control systems: A compositional approach. *Nonlinear Analysis: Hybrid Systems*, 43:101097, 2021.

[111] S. Liu, A. Saoud, P. Jagtap, D. V. Dimarogonas, and M. Zamani. Compositional synthesis of signal temporal logic tasks via assume-guarantee contracts. *arXiv preprint arXiv:2203.10041*, 2022.

[112] S. Liu, A. Swikir, and M. Zamani. Compositional verification of initial-state opacity for switched systems. In *59th IEEE Conference on Decision and Control (CDC)*, pages 2146–2151, 2020.

[113] S. Liu, A. Swikir, and M. Zamani. Verification of approximate opacity for switched systems: A compositional approach. *Nonlinear Analysis: Hybrid Systems*, 42:101084, 2021.

[114] S. Liu, A. Trivedi, X. Yin, and M. Zamani. Secure-by-construction synthesis of cyber-physical systems. *Annual Reviews in Control*, 53:30–50, 2022.

[115] S. Liu, X. Yin, and M. Zamani. On a notion of approximate opacity for discrete-time stochastic control systems. In *American Control Conference (ACC)*, pages 5413–5418. IEEE, 2020.

[116] S. Liu and M. Zamani. Compositional synthesis of almost maximally permissible safety controllers. In *American Control Conference (ACC)*, pages 1678–1683. IEEE, 2019.

[117] S. Liu and M. Zamani. Verification of approximate opacity via barrier certificates. *IEEE Control Systems Letters*, 5(4):1369–1374, 2020.

[118] S. Liu and M. Zamani. Compositional synthesis of opacity-preserving finite abstractions for interconnected systems. *Automatica*, 131:109745, 2021.

[119] A. Lomuscio and L. Maganti. An approach to reachability analysis for feed-forward relu neural networks. *arXiv preprint arXiv:1706.07351*, 2017.

[120] N. Lu, N. Cheng, N. Zhang, X. Shen, and J. W. Mark. Connected vehicles: Solutions and challenges. *IEEE Internet of Things Journal*, 1(4):289–299, 2014.

[121] Y. Lu and M. Zhu. Privacy preserving distributed optimization using homomorphic encryption. *Automatica*, 96:314–325, 2018.

[122] R. B. Lyngsø and C. N. Pedersen. The consensus string problem and the complexity of comparing hidden markov models. *Journal of Computer and System Sciences*, 65(3):545–569, 2002.

[123] C. Ma and W. Wonham. Nonblocking supervisory control of state tree structures. *IEEE Transactions on Automatic Control*, 51(5):782–793, 2006.

[124] K. Mai. *Side Channel Attacks and Countermeasures*, pages 175–194. Springer, New York, NY, 2012.

[125] R. Majumdar, K. Mallik, A.-K. Schmuck, and D. Zufferey. Assume–guarantee distributed synthesis. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 39(11):3215–3226, 2020.

[126] K. Mallik, A. Schmuck, S. Soudjani, and R. Majumdar. Compositional synthesis of finite-state abstractions. *IEEE Transactions on Automatic Control*, 64(6):2629–2636, 2018.

[127] P.-J. Meyer, A. Girard, and E. Witrant. Compositional abstraction and safety synthesis using overlapping symbolic models. *IEEE Transactions on Automatic Control*, 63(6):1835–1841, 2017.

[128] D. Milushev, W. Beck, and D. Clarke. Noninterference via symbolic execution. In *Formal Techniques for Distributed Systems*, pages 152–168. Springer, 2012.

[129] A. Mitra and S. Sundaram. Byzantine-resilient distributed observers for LTI systems. *Automatica*, 108:108487, 2019.

[130] Y. Mo and R. M. Murray. Privacy preserving average consensus. *IEEE Transactions on Automatic Control*, 62(2):753–765, 2017.

[131] S. Mohajerani, Y. Ji, and S. Lafortune. Compositional and abstraction-based approach for synthesis of edit functions for opacity enforcement. *IEEE Transactions on Automatic Control*, 2019.

[132] S. Mohajerani and S. Lafortune. Transforming opacity verification to nonblocking verification in modular systems. *IEEE Transactions on Automatic Control*, 65(4):1739–1746, 2019.

[133] A. Mohsen Nia, S. Sur-Kolay, A. Raghunathan, and N. K. Jha. Physiological information leakage: A new frontier in health information security. *IEEE Transactions on Emerging Topics in Computing*, 4(3):321–334, 2016.

[134] S. Nilizadeh, Y. Noller, and C. S. Păsăreanu. Diffuzz: differential fuzzing for side-channel analysis. In *41st International Conference on Software Engineering (ICSE)*, pages 176–187. IEEE, 2019.

[135] M. Noori-Hosseini, B. Lennartson, and C. Hadjicostis. Compositional visible bisimulation abstraction applied to opacity verification. *IFAC-PapersOnLine*, 51(7):434–441, 2018.

[136] M. Noori-Hosseini, B. Lennartson, and C. Hadjicostis. Incremental observer reduction applied to opacity verification and synthesis. arXiv:1812.08083, 2018.

[137] R. Oura, A. Sakakibara, and T. Ushio. Reinforcement learning of control policy for linear temporal logic specifications using limit-deterministic Büchi automata. *IEEE Control Systems Letters*, 4(3):761–766, 2020.

[138] A. Papachristodoulou, J. Anderson, G. Valmorbida, S. Prajna, P. Seiler, and P. Parrilo. SOSTOOLS version 3.00 sum of squares optimization toolbox for MATLAB. *arXiv preprint arXiv:1310.4716*, 2013.

[139] C. S. Pasareanu, Q.-S. Phan, and P. Malacaria. Multi-run side-channel analysis using symbolic execution and max-smt. In *29th Computer Security Foundations Symposium (CSF)*, pages 387–400. IEEE, 2016.

[140] A. Peruffo, D. Ahmed, and A. Abate. Automated formal synthesis of neural barrier certificates for dynamical models. *arXiv preprint arXiv:2007.03251*, 2020.

[141] A. Pneuli and R. Rosner. Distributed reactive systems are hard to synthesize. In *31st Annual Symposium on Foundations of Computer Science*, volume 2, pages 746–757, 1990.

[142] G. Pola, E. De Santis, and M. Di Benedetto. Approximate diagnosis of metric systems. *IEEE Control Systems Letters*, 2(1):115–120, 2018.

[143] G. Pola, E. De Santis, M. D. Di Benedetto, and D. Pezzuti. Design of decentralized critical observers for networks of finite state machines: A formal method approach. *Automatica*, 86:174–182, 2017.

[144] G. Pola and M. D. Di Benedetto. Control of cyber-physical-systems with logic specifications: A formal methods approach. *Annual Reviews in Control*, 47:178–192, 2019.

[145] G. Pola, A. Girard, and P. Tabuada. Approximately bisimilar symbolic models for nonlinear control systems. *Automatica*, 44(10):2508–2516, 2008.

[146] G. Pola, P. Pepe, and M. Di Benedetto. Symbolic models for networks of control systems. *IEEE Transactions on Automatic Control*, 61(11):3663–3668, 2016.

[147] G. Pola and P. Tabuada. Symbolic models for nonlinear control systems: Alternating approximate bisimulations. *SIAM Journal on Control and Optimization*, 48(2):719–733, 2009.

[148] S. Prajna and A. Jadbabaie. Safety verification of hybrid systems using barrier certificates. In *International Workshop on Hybrid Systems: Computation and Control (HSCC)*, pages 477–492. Springer, 2004.

[149] S. Prajna, A. Jadbabaie, and G. J. Pappas. A framework for worst-case and stochastic safety verification using barrier certificates. *IEEE Transactions on Automatic Control*, 52(8):1415–1428, 2007.

[150] S. Prajna and A. Rantzer. Primal–dual tests for safety and reachability. In *International Workshop on Hybrid Systems: Computation and Control (HSCC)*, pages 542–556. Springer, 2005.

[151] L. Pulina and A. Tacchella. Challenging smt solvers to verify neural networks. *Ai Communications*, 25(2):117–135, 2012.

[152] A. Raghunathan and N. K. Jha. Hijacking an insulin pump: Security attacks and defenses for a diabetes therapy system. In *13th International Conference on e-Health Networking, Applications and Services*, pages 150–156, 2011.

[153] B. Ramasubramanian, R. Cleaveland, and S. Marcus. A framework for decentralized opacity in linear systems. In *54th Annual Allerton Conference on Communication, Control, and Computing*, pages 274–280, 2016.

[154] B. Ramasubramanian, R. Cleaveland, and S. Marcus. A framework for opacity in linear systems. In *American Control Conference*, pages 6337–6344, 2016.

[155] B. Ramasubramanian, R. Cleaveland, and S. Marcus. Opacity for switched linear systems: Notions and characterization. In *56th IEEE Conference on Decision and Control*, pages 5310–5315, 2017.

[156] J. H. Reif. The complexity of two-player games of incomplete information. *Journal of Computer and System Sciences*, 29(2):274–301, 1984.

[157] G. Reissig, A. Weber, and M. Rungger. Feedback refinement relations for the synthesis of symbolic controllers. *IEEE Transactions on Automatic Control*, 62(4):1781–1796, 2017.

[158] B. S. Ruffer. Monotone dynamical systems, graphs, and stability of largescale interconnected systems. *Ph.D. thesis, Fachbereich 3, Mathematik und Informatik, Universität Bremen, Germany*, 2007.

[159] M. Rungger and M. Zamani. Compositional construction of approximate abstractions of interconnected control systems. *IEEE Transactions on Control of Network Systems*, 5(1):116–127, 2016.

[160] A. Saboori and C. Hadjicostis. Notions of security and opacity in discrete event systems. In *2007 46th IEEE Conference on Decision and Control*, pages 5056–5061, 2007.

[161] A. Saboori and C. Hadjicostis. Verification of k-step opacity and analysis of its complexity. *IEEE Transactions on Automation Science and Engineering*, 8(3):549–559, 2011.

[162] A. Saboori and C. Hadjicostis. Verification of initial-state opacity in security applications of discrete event systems. *Inform. Sci.*, 246:115–132, 2013.

[163] A. Saboori and C. Hadjicostis. Current-state opacity formulations in probabilistic finite automata. *IEEE Transactions on Automatic Control*, 59(1):120–133, 2014.

[164] A. Saboori and C. N. Hadjicostis. Reduced-complexity verification for initial-state opacity in modular discrete event systems. *IFAC Proceedings Volumes*, 43(12):78–83, 2010.

[165] A. Saboori and C. N. Hadjicostis. Verification of infinite-step opacity and complexity considerations. *IEEE Transactions on Automatic Control*, 57(5):1265–1269, 2012.

[166] Y. E. Sahin, N. Ozay, and S. Tripakis. Multi-agent coordination subject to counting constraints: A hierarchical approach. In *Distributed Autonomous Robotic Systems*, pages 265–281. Springer, 2019.

[167] H. Sandberg, S. Amin, and K. Johansson. Cyberphysical security in networked control systems. *IEEE Control Systems*, 35(1):20–23, 2015.

[168] C. Santoyo, M. Dutreix, and S. Coogan. A barrier function approach to finite-time stochastic system verification and control. *Automatica*, 125:109439, 2021.

[169] A. Saoud, A. Girard, and L. Fribourg. Assume-guarantee contracts for continuous-time systems. *Automatica*, 134:109910, 2021.

[170] S. Schewe. *Synthesis of distributed systems*. PhD thesis, Saarland University, Saarbrücken, Germany, 2008.

[171] S. Schewe. Distributed synthesis is simply undecidable. *Information Processing Letters*, 114(4):203 – 207, 2014.

[172] S. Schewe and B. Finkbeiner. Bounded synthesis. In *International Symposium on Automated Technology for Verification and Analysis*, pages 474–488, 2007.

[173] P. Schillinger, M. Bürger, and D. V. Dimarogonas. Simultaneous task allocation and planning for temporal logic goals in heterogeneous multi-robot systems. *The international journal of robotics research*, 37(7):818–838, 2018.

[174] S. Schinzel. An efficient mitigation method for timing side channels on the web. In *2nd International Workshop on Constructive Side-Channel Analysis and Secure Design (COSADE)*, 2011.

[175] C. E. Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[176] M. Sharf, B. Besselink, A. Molin, Q. Zhao, and K. H. Johansson. Assume/guarantee contracts for dynamical systems: Theory and computational tools. *IFAC-PapersOnLine*, 54(5):25–30, 2021.

[177] G. Smith. On the foundations of quantitative information flow. In *International Conference on Foundations of Software Science and Computational Structures*, pages 288–302. Springer, 2009.

[178] M. Sousa and I. Dillig. Cartesian hoare logic for verifying k-safety properties. In *37th ACM SIGPLAN Conference on Programming Language Design and Implementation*, volume 51, pages 57–69, 2016.

[179] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization methods and software*, 11(1-4):625–653, 1999.

[180] 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.

[181] R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[182] A. Swikir, A. Girard, and M. Zamani. From dissipativity theory to compositional synthesis of symbolic models. In *Indian Control Conference (ICC)*, pages 30–35. IEEE, 2018.

[183] A. Swikir and M. Zamani. Compositional abstractions of interconnected discrete-time switched systems. In *18th European Control Conference*, pages 1251–1256, 2019.

[184] A. Swikir and M. Zamani. Compositional synthesis of finite abstractions for networks of systems: A small-gain approach. *Automatica*, 107:551–561, 2019.

[185] A. Swikir and M. Zamani. Compositional synthesis of symbolic models for networks of switched systems. *IEEE Control Systems Letters*, 3(4):1056–1061, 2019.

[186] P. Tabuada. *Verification and Control of Hybrid Systems: A Symbolic Approach*. Springer Science & Business Media, 2009.

[187] A. Tarski. A decision method for elementary algebra and geometry. In *Quantifier elimination and cylindrical algebraic decomposition*, pages 24–84. Springer, 1998.

[188] Y. Tazaki and J. Imura. Bisimilar finite abstractions of interconnected systems. In M. Egerstedt and B. Mishra, editors, *International Conference on Hybrid Systems: Computation and Control (HSCC)*, volume 4981, pages 514–527. Springer Verlag, Berlin Heidelberg, 2008.

[189] S. Tizpaz-Niari, P. Cerný, and A. Trivedi. Quantitative mitigation of timing side channels. In *International Conference on Computer Aided Verification (CAV)*, pages 140–160, 2019.

[190] Y. Tong and H. Lan. Current-state opacity verification in modular discrete event systems. In *58th IEEE Conference on Decision and Control (CDC)*, pages 7665–7670, 2019.

[191] Y. Tong, Z. Li, C. Seatzu, and A. Giua. Decidability of opacity verification problems in labeled Petri net systems. *Automatica*, 80:48–53, 2017.

[192] 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.

[193] D. N. Tran. *Advances in stability analysis for nonlinear discrete-time dynamical systems*. PhD thesis, Univ. Newcastle, 2018.

[194] J. Tumova and D. V. Dimarogonas. Multi-agent planning under local LTL specifications and event-based synchronization. *Automatica*, 70:239–248, 2016.

[195] L. Vu, D. Chatterjee, and D. Liberzon. Input-to-state stability of switched systems and switching adaptive control. *Automatica*, 43(4):639–646, 2007.

[196] S. Walters. How can drones be hacked? `https://medium.com/@swalters/how-can-drones-be-hacked-the-updated-list-of-vulnerable-drones-attack-tools-dd2e006d6809`, 2016. Online published 19-Oct-2016.

[197] L. Wang, A. D. Ames, and M. Egerstedt. Safety barrier certificates for collisions-free multirobot systems. *IEEE Transactions on Robotics*, 33(3):661–674, 2017.

[198] T. Wongpiromsarn, U. Topcu, and A. Lamperski. Automata theory meets barrier certificates: Temporal logic verification of nonlinear systems. *IEEE Transactions on Automatic Control*, 61(11):3344–3355, 2015.

[199] B. Wu and H. Lin. Privacy verification and enforcement via belief abstraction. *IEEE Control Sys. Letters*, 2(4):815–820, 2018.

[200] B. Wu, Z. Liu, and H. Lin. Parameter and insertion function co-synthesis for opacity enhancement in parametric stochastic discrete event systems. In *American Control Conference*, pages 3032–3037, 2018.

[201] M. Wu, S. Guo, P. Schaumont, and C. Wang. Eliminating timing side-channel leaks using program repair. In *27th ACM SIGSOFT International Symposium on Software Testing and Analysis*, pages 15–26, 2018.

[202] Y.-C. Wu and S. Lafortune. Comparative analysis of related notions of opacity in centralized and coordinated architectures. *Discrete Event Dynamic Systems*, 23(3):307–339, 2013.

[203] W. Xiang and T. T. Johnson. Reachability analysis and safety verification for neural network control systems. *arXiv preprint arXiv:1805.09944*, 2018.

[204] J. Yang, W. Deng, and D. Qiu. Current-state opacity and initial-state opacity of modular discrete event systems. *International Journal of Control*, pages 1–24, 2021.

[205] J. Yang, W. Deng, D. Qiu, and C. Jiang. Opacity of networked discrete event systems. *Information Sciences*, 543:328–344, 2021.

[206] S. Yang, J. Hou, X. Yin, and S. Li. Opacity of networked supervisory control systems over insecure communication channels. *IEEE Transactions on Control of Network Systems*, 8(2):884–896, 2021.

[207] S. Yang and X. Yin. Secure your intention: On notions of pre-opacity in discrete-event systems. *arXiv preprint arXiv:2010.14120*, 2020.

[208] K. Yazdani, A. Jones, K. Leahy, and M. Hale. Differentially private LQ control. *arXiv preprint arXiv:1807.05082*, 2018.

[209] 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.

[210] X. Yin and S. Li. Verification of opacity in networked supervisory control systems with insecure control channels. In *57th IEEE Conference on Decision and Control (CDC)*, pages 4851–4856, 2018.

[211] 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.

[212] X. Yin, M. Zamani, and S. Liu. On approximate opacity of cyber-physical systems. *IEEE Transactions on Automatic Control*, 66(4):1630–1645, 2021.

[213] X. Yu and Y. Xue. Smart grids: A cyber–physical systems perspective. *Proceedings of the IEEE*, 104(5):1058–1070, 2016.

[214] M. Zamani, A. Abate, and A. Girard. Symbolic models for stochastic switched systems: A discretization and a discretization-free approach. *Automatica*, 55:183–196, 2015.

[215] M. Zamani and M. Arcak. Compositional abstraction for networks of control systems: A dissipativity approach. *IEEE Transactions on Control of Network Systems*, 5(3):1003–1015, 2018.

[216] M. Zamani, P. M. Esfahani, R. Majumdar, A. Abate, and J. Lygeros. Symbolic control of stochastic systems via approximately bisimilar finite abstractions. *IEEE Transactions on Automatic Control*, 59(12):3135–3150, 2014.

[217] M. Zamani, M. Mazo, M. Khaled, and A. Abate. Symbolic abstractions of networked control systems. *IEEE Transactions on Control of Network Systems*, 5(4):1622–1634, 2018.

[218] M. Zamani, G. Pola, M. Mazo, and P. Tabuada. Symbolic models for nonlinear control systems without stability assumptions. *IEEE Transactions on Automatic Control*, 57(7):1804–1809, 2012.

[219] G. Zames. On the input-output stability of time-varying nonlinear feedback systems part one: Conditions derived using concepts of loop gain, conicity, and positivity. *IEEE transactions on automatic control*, 11(2):228–238, 1966.

[220] D. Zhang, A. Askarov, and A. C. Myers. Predictive mitigation of timing channels in interactive systems. In *Proceedings of the 18th ACM conference on Computer and communications security*, pages 563–574. ACM, 2011.

[221] D. Zhang, A. Askarov, and A. C. Myers. Language-based control and mitigation of timing channels. *SIGPLAN Notices*, 47(6):99–110, 2012.

[222] K. Zhang, X. Yin, and M. Zamani. Opacity of nondeterministic transition systems: A (bi)simulation relation approach. *IEEE Transactions on Automatic Control*, 64(12):5116–5123, 2019.

[223] Z. Zhang, S. Shu, and C. Xia. Networked opacity for finite state machine with bounded communication delays. *Information Sciences*, 572:57–66, 2021.

[224] G. Zinck, L. Ricker, H. Marchand, and L. Hélouët. Enforcing opacity in modular systems. *IFAC-PapersOnLine*, 53(2):2157–2164, 2020.