Analysis and prevention of stealthy aging attacks : an approach based on dynamical systems theory

Cédric Escudero

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE GRENOBLE ALPES

Spécialité : AUTOMATIQUE - PRODUCTIQUE

Arrêté ministériel : 25 mai 2016

Présentée par

Cédric ESCUDERO

Thèse dirigée par **Eric ZAMAI**, Professeur des Universités
et codirigée par **Paolo MASSIONI**, Maître de Conférences

préparée au sein du **Laboratoire des Sciences pour la Conception, l'Optimisation et la Production de Grenoble**
dans **l'École Doctorale Electronique, Electrotechnique, Automatique, Traitement du Signal (EEATS)**

Analyse et prévention des attaques cachées de vieillissement : une approche basée sur la théorie des systèmes dynamiques

Analysis and prevention of stealthy aging attacks: an approach based on dynamical systems theory

Thèse soutenue publiquement le **15 janvier 2021,**
devant le jury composé de :

Monsieur Jean-Marc THIRIET
PROFESSEUR DES UNIVERSITES, Université Grenoble Alpes, Président du jury
Madame Mireille BAYART MERCHEZ
PROFESSEURE DES UNIVERSITES, Université de Lille, Rapporteur
Monsieur Frédéric KRATZ
PROFESSEUR DES UNIVERSITES, INSA Centre Val de Loire, Rapporteur
Monsieur Carlos MURGUIA RENDON
PROFESSEUR ASSISTANT, Eindhoven University of Technology, Examinateur
Monsieur Stéphane DE FLEURIAU
INGENIEUR, Direction Générale de l'Armement, Invité
Monsieur Franck SICARD
DOCTEUR INGENIEUR, Naval Group, Invité
Monsieur Paolo MASSIONI
MAITRE DE CONFERENCES, INSA de Lyon, Co-directeur de thèse
Monsieur Eric ZAMAI
PROFESSEUR DES UNIVERSITES, INSA de Lyon, Directeur de thèse

Acknowledgments

The research works presented in this manuscript have been done at Laboratoire des Sciences pour la Conception, l'Optimisation et la Production (G-SCOP) in the research team Gestion et Conduite des Systèmes de Production (GCSP). That is why, I would like to thank M. François Villeneuve, G-SCOP Research Laboratory Director, for having welcomed and allowed me to conduct my research works during these three years. Also, I would like to thank M. Bruno Allard and M. Eric Bideaux, respectively Director of Ampère Research Laboratory and Head of the Department Méthodes pour l'Ingénierie des Systèmes (MIS) at Ampère Research Laboratory, for having welcomed me as an invited researcher for my last year.

First and foremost, I would like to express my gratitude to my PhD supervisor, M. Eric Zamaï, Full Professor at Institut National des Sciences Appliquées Lyon (INSA Lyon), for having accepted to supervise me. In addition to this, I would like to thank him for the quality of his supervision, his availability and his advice.

I would like also to express my gratitude to my PhD co-supervisor, M. Paolo Massioni, Associate Professor at INSA Lyon, for having accepted me to supervise me. In addition to this, I would like to thank him for the quality of his supervision and his advice.

I would like to thank M. Jean-Marc Thiriet, Full Professor at Université Grenoble Alpes (UGA), for having accepted to preside the jury of my PhD Defense.

I would like to thank Ms. Mireille Bayart Merchez, Full Professor at Université de Lille, and M. Frédéric Kratz, Full Professor at INSA Centre Val de Loire, for the interest they have shown on my research works and for having accepted to be reviewers of my research works.

I would like to thank M. Carlos Murguia Rendon, Assistant Professor at Eindhoven University of Technology, for having accepted to examine my research works and to participate to my PhD Defense.

I would like to thank M. Stéphane De Fleuriau, Engineer at Direction Générale de l'Armement (DGA), and M. Franck Sicard, PhD-Engineer at Naval Group, for having accepted to participate to my PhD Defense.

I would like to thank M. Nouredine Hadjsaid, Director of Laboratoire de Génie Electrique de Grenoble (G2ELab), and M. Jean-Paul Jamont, Associate Professor at UGA, for having accepted to follow the progress of my PhD.

I would like also to thank the collaborators I worked with during these three years: M. Bertrand Raison, Full Professor at UGA, M. Quoc Bao Duong, research engineer at S.mart Grenoble-Alpes, M. Laurent Thibault, Lecturer at UGA, M. Romain Delpoux, Associate Professor at INSA Lyon, M. Vincent Léchappé, Associate Professor at INSA

Lyon, M. Sébastien Henry, Associate Professor at IUT Lyon 1.

I am also grateful to my close friends M. Iliasse Benbalaid and M. Lucas Mourey for the time spent to read my papers.

I would like to thank my beloved family and in particular my mother, my father, my sister, my brother and my grand mother for having supported me. I would like also to dedicate this thesis to my grand fathers.

Lastly, I would like to thank all my friends and in particular the ones with who I spent most of my time during these three years: M. Akram Chergui, M. Mouhamadou Mansour Mbow, and Ms. Cléa Martinez.

Contents

List of Figures

List of Tables

List of Acronyms

List of Abbreviations

List of Publications

Journals:

- **C. Escudero**, F. Sicard, A. Beaudet and E. Zamaï, E. *Security of the Process Integrity under Attacks Propagation in Industrial Control Systems: a Classification of Control-based Methods. Reliability Engineering and System Safety (Reliab. Eng. Syst. Saf.)*, **(Submitted)**, pages 1-1, Sept. 2020

- **C. Escudero**, P. Massioni, E. Zamaï and B. Raison. *Control System Theory Methods to Analyze, Prevent and Design Stealthy Aging Attacks. IEEE Transaction on Industrial Informatics (IEEE Trans. Ind. Informat.)*, **(Submitted)**, pages 1-1, Sept. 2020

Conferences:

- A. Beaudet, **C. Escudero** and E. Zamaï. *Malicious Anomaly Detection Approaches Robustness in Manufacturing ICSs.* In 2021 IFAC 17th Symposium on Information Control Problems in Manufacturing (INCOM), **(Submitted)**, pages 1-8, Jun. 2021

- **C. Escudero**, P. Massioni, G. Scorletti and E. Zamaï. *Security of Control Systems: Prevention of Aging Attacks by means of Convex Robust Simulation Forecasts.* In 2020 IFAC World Congress, pages 1-8, Jul. 2020

- E. M. Merouane, **C. Escudero**, F. Sicard and E. Zamaï. *Aging Attacks against Electro-Mechanical Actuators from Control Signal Manipulation.* In 2020 IEEE International Conference on Industrial Technology (ICIT), pages 133-138, Feb. 2020

- A. Beaudet, F. Sicard, **C. Escudero** and E. Zamaï. *Process-Aware Model-based Intrusion Detection System on Filtering Approach: Further Investigations.* In 2020 IEEE International Conference on Industrial Technology (ICIT), pages 310-315, Feb. 2020

- **C. Escudero**, and E. Zamaï. *Prevention of Aging Attacks: Malicious Nature of the Control Signal.* In 2019 IEEE International Automatic Control Conference (CACS), pages 1-6, Nov. 2019

- F. Sicard, **C. Escudero**, E. Zamaï and J.-M. Flaus. *From ICS Attacks' Analysis to the S.A.F.E. Approach: implementation of Filters based on behavioral models and critical state distance for ICS cybersecurity.* In 2018 IEEE 2nd Cyber Security in Networking Conference (CSNet), pages 1-8, Oct. 2018

- **C. Escudero**, F. Sicard and E. Zamaï. *Process-Aware Model based IDSs for Industrial Control Systems Cybersecurity: Approaches, Limits and Further Research.* In 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), volume 1, pages 605-612, Sep. 2018

Workshops:

- **C. Escudero** and E. Zamaï. *Sécurité des Systèmes Déterministes Linéaires: Composantes Malveillantes du Signal de Commande des Actionneurs.* Extended abstract for the National Research Group Workshop Modélisation, Analyse et Conduite des Systèmes dynamiques (GdR MACS), Jun. 2019

Introduction

Introduction

Industrial Control System (ICS) are architectures controlling a physical system to achieve an industrial objective. They are present in various sectors including energy generation and distribution, water treatment, manufacturing production, aerospace and defense. Those architectures are equipped with various digital entities, including Human Machine Interface (HMI), Programmable Logic Controller (PLC), regulator, organized for reducing the complexity of the control. Historically, ICSs have been designed to improve the productivity, but the cybersecurity has not been considered. Due to this lack, ICSs are facing cyberattacks. Plenty of them manipulating the architecture have been reported in the literature. They aim to steal sensitive information or to violate the integrity of the physical system. The violation of the physical system integrity refers to an intended alteration or destruction of the physical system through its control. It leads to a partial or complete failure of the services delivered by the physical system. Hence, cyberattacks are a new root-cause of failure, that we call the malicious acts. Those malicious acts aim to create and propagate anomalies in the architecture by exploiting vulnerabilities of the digital entities.

In this thesis, we address a new type of attack that aims to accelerate the aging of the actuators. The attacker is considered to manipulate the control signal sent to them. The objective of this thesis is to develop methods for analyzing such attacks and preventing them. The analysis consists in quantifying the potential impact a control signal manipulation could have on the process. The prevention consists in restraining the control signal such that stealthy aging attacks cannot occur.

This thesis addresses the following scientific issues:

- Modeling the attack stealthiness in set theory,
- Stability condition required from the invariant set theory,
- Loss of the temporal variable in the invariant set theory.

This manuscript is organized in Part as follows.

Part 1 presents the general problematic of our research works. After having presented the general context about the malicious act acting on ICSs to reach the process integrity, we position our research works in the controllers. In particular, we focus on the continuous controllers. Then, the problem of anomaly observation is explained. It appears that the process integrity can be violated if the control signal is abnormal. Thus, we orientate our works on anomalies in the control signal. In particular, our interest concerns anomalies accelerating the aging of the actuators. After having briefly reviewed the literature, we position our works on the set-theoretic methods and linear optimization.

Part 2 provides our contributions. First, the analysis and prevention of stealthy aging attacks are restrained to a subcase, denoted Attack Type 1 (AT1). AT1 consists

in manipulating the control signal to let the dynamical system reach a subset of the state space that features an aging faster than normal. We propose two set-theoretic methods for quantifying the impact of AT1 on the process. The first method based on the invariant notion quantifies it on the infinite-time horizon. The second one based on robust simulation quantifies it on the finite-time horizon. A set of Algorithms is proposed for both methods to be applicable on a dynamical system. Lastly, a method based on linear optimization is proposed for assessing the potential impact stealthy aging attack of a more general case could lead on the process. This more general case Attack Type 2 (AT2) consists in maximizing the time integral of a degradation function modeling the stresses in the system.

Part 3 proposes to apply the methods upon two application examples: a thin film manufacturing system, and a satellite system. The results are analyzed and show that the robust simulation-based method is more adapted to deal with stealthy aging attacks rather than the invariant-set based one. In addition to this, we show that stealthy aging attacks seem more dangerous on over actuated dynamical systems as the abnormal control actions can be compensated by the other actuators to hide the attack regarding the delivered service.

Part I

General framework

Chapter 1

Malevolence: a new root cause of failures in Industrial Control Systems

Contents

1.1 Introduction

This Chapter presents the overall context of this thesis. It takes place in ICSs, where heterogeneous industrial equipment and devices operate each other to control a physical system. After having presented what is an ICS in Section 1.2, the problem of faults in ICS is explained and the response given by the community is detailed in Section 1.3. This response has introduced a new problem in ICSs: the malevolence. In this thesis, we focus on the malevolence that aim to violate the integrity of the physical system. That is why, the malevolence is studied in Section 1.4 with the viewpoint of the integrity of the physical system.

1.2 Industrial Control Systems

Many sectors including critical ones [Rushby 1994] have quickly integrated automation of their industrial operations such as energy generation and distribution, water treatment, manufacturing production, aerospace and defense [Nozaic & Freese 2009, Stouffer *et al.* 2015, Groover 2008]. This automation raise stems from the fact companies are constantly chasing better performance. In the following subsection, a brief study of ICSs is given.

1.2.1 General description

ICSs are organized around workstations, mostly automated. The workstations perform industrial operations using industrial equipment, i.e. the physical system used to realize the industrial operations. These workstations are interconnected with a control system devoted to the coordination of the control, and to the control of the workstations. The automated operations are often supported by manual interventions of operators. These manual interventions are dedicated to auxiliary operations such as tools changing or the loading and unloading of workpieces in manufacturing productions, cleaning of screens in water treatment systems, or repair and replace activities.
ICSs aim at:

- Relieving the human of the tedious and dangerous tasks (e.g. operations in environment with high temperature or toxic gas) [Robla-Gómez *et al.* 2017, Cherubini *et al.* 2016],
- Performing complex industrial operations, almost impossible by the human (e.g. manufacturing precision in aerospace and defense),
- Improving the productivity (e.g. production time, delivery times),
- Reducing the production cost,
- Optimizing energetic performance during the industrial operations.

To successfully achieve its objectives, an ICS requires flexibility [Michalos *et al.* 2010]. The flexibility is the ability to easily adapt to changes. Industrial equipment flexibility allows flexibility in the operations realizable by the equipment. This requirement plays a major role to deal with failures in ICSs. However, the industrial equipment flexibility is not enough. In complementary, flexibility in the control system is required to easily adapt to industrial objectives and failures. Both requirements allow the industrial equipment and the control system to respond to industrial objectives changes, and failures in the ICS [Zamai *et al.* 1998].

1.2.2 Internal structure

In general, an ICS has an internal structure split into three parts as illustrated in Figure 1.1:

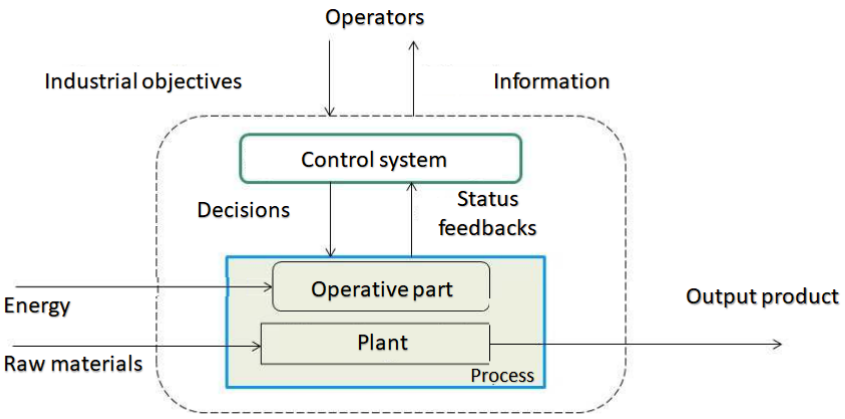


Figure 1.1: General structure of ICS

- **The plant** represents the physical entities which are transformed to achieve the industrial objectives such as displacement of a workpiece, water level in a tank, flow of power in power lines. These physical entities are often called raw materials and represent respectively the workpiece, the water and the power in the previous examples.

- **The operative part** is the set of actuators and sensors, also called field devices, or industrial control and instrumentation devices. On the one hand, actuators receive decisions from the control system and transform them into physical actuation (e.g. angular velocity, electrical current, electromagnetic torque, force, fluid volumetric rate of flow). Physical actuation is applied on the plant to transform the physical entities. To achieve this transformation, actuators use energy: electrical, pneumatic, or hydraulic. On other hand, sensors measure physical quantities related to the entities transformation. These measurements are transmitted back to the control system as status feedbacks. The operative part together with the plant is usually gathered into a single term: the **process**, being the physical system to be controlled.

- **The control system** coordinates the control of the process by transmitting decisions to the actuators in order to impose a desired behavior to the plant based on the status feedbacks and the industrial objectives. The control system can transmit relevant information about the process, or even itself to operators. The operators interconnected with the control system monitor the control of the plant, and can make decisions if required.

1.2.3 Control system

As seen previously, the control system is the core of an ICS as it coordinates the control and controls the process. However, the plant involves plenty of physical phenomenon which need to be considered by the control system in order to impose the desired behavior. Each physical phenomenon needs to be controlled through actuators, and monitored

through sensors, thus the higher the number of physical phenomenon, the higher the number of actuators and sensors is. This complexity has increased in the past with more complex process to be controlled. From this raise, it quickly appeared that centralized control system was not anymore adapted. This major problem has been considered in the past, and many architectures have emerged to deal with [Jones & Saleh 1989]. One of the most used architecture is the hierarchical and modular control system architecture, known under the name "Computer Integrated Manufacturing" [Williams 1990] and normalized nowadays in the ISA95 [ISA]. Although each sector has specific features, their architecture follows this conceptual architecture. It splits the control system into five levels of lower complexities as illustrated in Figure 1.2:

- Level 5: Management information
- Level 4: Production scheduling and operational management
- Level 3: Supervisory control
- Level 2: Control coordinator
- Level 1: Local control

Levels 5 and 4 are dedicated to the production scheduling and management information, while the levels 3, 2, and 1 are devoted to the process control. These levels are interconnected together through communication levels. Throughout this thesis, only the levels devoted to the process control are of concern, i.e. levels 3, 2, and 1.

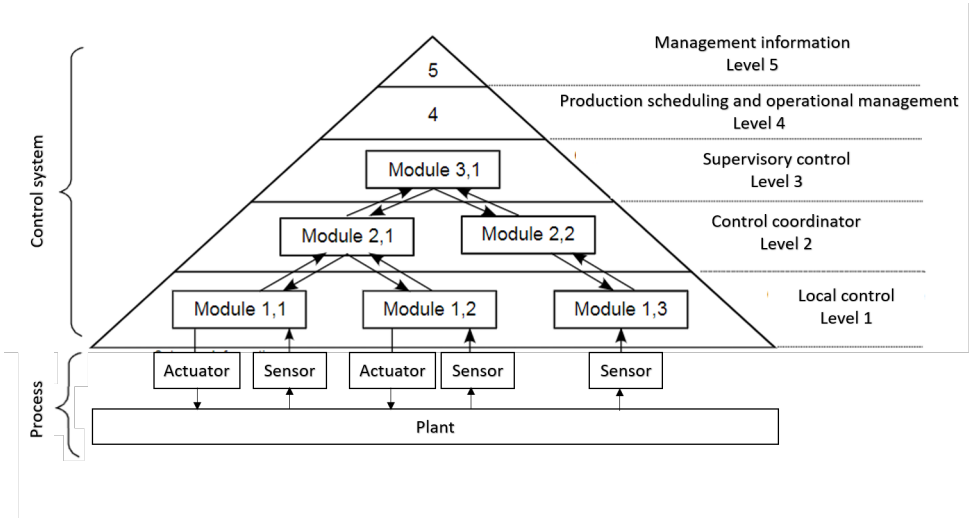


Figure 1.2: Standard ICS architecture

Each level has several modules, and each module is a digital device having a control mission in the control system, as it will be detailed in Chapter 2. The operating principle of this control system architecture is based on a request-reply procedure [Jones & Saleh 1989], illustrated in Figure 1.3 and explained as follows. A module at a level n transmits a decision to a module at a level $n-1$. The latter module breaks down the

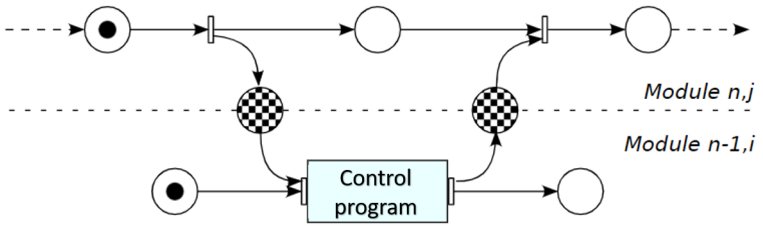


Figure 1.3: Request-reply procedure

decision into m other decisions, and transmits them to the lower layer, and so on until the decisions reach the process.

After the process has reacted to these decisions, m status feedbacks, corresponding to the reaction response of the process to the m decisions, are propagated to the modules. These status feedbacks inform about the activity of the process. They can be either normal, or abnormal if the status feedback does not correspond to the expected one, i.e. expected activity, [Duong *et al.* 2013]. In the last case, this anomaly is commonly known as a fault. Since ICSs are made of many industrial equipment ranging from dozens to hundreds of equipment, faults issue is a complex problem to deal with.

1.3 Responsiveness to faults

As this thesis concerns attacks leading to a fault in the operations of the ICS, faults in ICS are presented in this section. First of all, the concept of failure commonly agreed in the literature is presented.

1.3.1 Failures in an ICS

A failure of an entity is defined as the cessation of its ability to deliver its service. A failure can lead to a fault in an entity. A fault of an entity is defined as its inability to deliver the service [Villemeur 1992]. A fault results in a deviation from a characteristic property or a property of the service to deliver [Isermann 1997][ISO13849-1 2015]. This deviation is seen as an anomaly in the ICS. A failure can result from the occurrence of other failures. The first failure is the origin of failure.

The origin of failures are usually classified into two groups: external or internal, of the architecture. The external failures can come from:

- The production scheduling and management information, i.e. levels 5 and 4, (e.g. industrial objectives changes, production rate changes).
- The raw materials (e.g. not fit with required specifications).
- The environment of the ICS (e.g. operators interacting with the ICS, energy used by actuators).

The internal failures can come from:

- The digital entities of the ICS, i.e. digital devices and digital communication, (e.g. failure of a communication, design error in a control program).
- The process: the failure of a sensor, of an actuator, or of the plant.

1.3.2 Root causes of failures

In ICSs, the root causes of failures are often classified into five main root causes: Man, Equipment, Method and Recipe, Material, and Environment [Ishikawa 1990, Doggett 2004]. The main root causes are explained below only with respect to facts related to the process control [Nguyen *et al.* 2016], as this is the interest of this thesis. Hence, the Man root cause refers to the exploitation and maintenance of equipment, and the Equipment refers to an intrinsic deviation from its technical specifications. In addition, as Method devotes to the recognition at work, so only the Recipe is studied. It concerns their characterization as this step is done without considering the current operations in the ICS, i.e. without the current stresses the equipment is subject to. The Material refers to the non-compliance with required specifications. Finally, the Environment is not considered as it devotes to the well-being at work.

- **Man** is, on the one hand, an irreplaceable stakeholder with adaptation and responsiveness capabilities to deal with unpredictable events. It is particularly essential in ICS with high variability such as electric power systems, or some manufacturing production systems (e.g. automotive assembly lines [Michalos *et al.* 2010], semiconductor manufacturing [Bouaziz 2012]). On the other hand, it can generate disturbances because of whether a lack of information to make its decisions for responding to a failure, or it is fallible in the tasks it realizes (e.g. negligence in the maintenance operations, tension variations causing paper web breaks due to control design error [Hristopulos & Uesaka 2002]) [Kobbacy & Murthy 2008, Proctor & Zandt 2018].
- **Equipment** is subject to an intrinsic aging process due to the various stresses it goes through during its control. The aging process comes from the phenomenon of wear and tear of the equipment's components [Dhillon 1999, Viveros *et al.* 2014]. The failure of an equipment can be predicted by statistical and probabilistic methods [Vrignat *et al.* 2015]. However, these methods usually consider normal conditions of equipment operation, often not satisfied in the ICS. Indeed, non-predicted failure of equipment's components have been reported owing to abnormal operations. Furthermore, [Isermann 2006] highlights the phenomenon of rapid technological change, and the consequences of the lack of knowledge about the new equipment causing an increase in the number of failures.
- **Recipe** is usually developed in the R&D department [Erickson & Hedrick 1999], and defined for dedicated equipment. A recipe specifies the industrial operations such as the equipment to be used, their setting parameters, and the starting and ending date of operations. A set of recipes is usually associated with a Work Order (WO) [Muñoz *et al.* 2011]. A WO is a set of orders the control coordinator level in the ICS architecture has to execute. A WO corresponds to a set of decisions

the operators have made in the supervisory control level as presented in Section 1.2. The recipes are usually developed in a dedicated environment without considering the stresses the equipment are subject to in current operations. Indeed, sometimes the ICS equipment operates beyond its operating conditions, and then operates in stress conditions. Therefore, the recipes can result in failures once they are executed by the control coordinator level [Muñoz *et al.* 2011]. In addition, the higher the stressed equipment, the higher the number of failures can occur [Hubac & Zamaï 2013].

- **Material** is the raw materials introduced in the ICS to realize industrial operations on it, as mentioned in Section 1.2. If the raw materials do not conform to required specification standards, then they can produce failures on the physical entities to transform, i.e. the plant, (e.g. web breakage due to low paper properties in newsprint production [Haapala *et al.* 2010]) [Ahmad *et al.* 2018], and even on the equipment transforming the raw materials (e.g. reactor accidents in chemical process industry [Kidam & Hurme 2013]).

1.3.3 Responsiveness of the control system

As faults are many in an ICS, it appeared the control systems were not designed to deal with. Indeed, the request-reply procedure was designed to optimize the operations in ICS by transmitting a decision, and waiting for its realization before transmitting another decision. However, it was not designed to deal with faults. The lack of response to faults in an ICS has appeared as the main cause of lack of performance in an ICS. As a result, ICSs have progressively evolved to improve its responsiveness, as detailed below.

In 1980, a first observation was stated: the control system and the industrial control and instrumentation devices, i.e. level 1 and 0 of the ICS, communicate in a single-way analog communication with point-to-point connections. Besides the high installation and maintenance cost it leads to, i.e. a lot of electrical wirings, the need for a two-way digital communication appeared in order to improve the feedback of information for fault-detection purposes. Indeed, the analog communications limited the feedback of information from the equipment to the control system in different aspects: the amount, the quality, and the cost of supplementary electrical wiring for the two-way communication. This conducts the progressive replacement of the analog communication by a serial digital communication, called fieldbus, so that more information can flow in both directions between the field devices and the control system. The emergence of integrated circuit in the same years has reinforced the need for digital processing capabilities in the actuators and sensors, so that they can transmit more information. These new digital processing capabilities and the emergence of fieldbuses have conducted progressively the replacement of analog devices to digital ones [Thomesse 2005, Galloway & Hancke 2013].

In 2000, a second observation was stated: the different levels of the ICS communicate with incompatible networking concepts, i.e. fieldbuses, and mostly Ethernet, and Internet Protocol (IP)-based Local Area Network (LAN) for the higher levels-. Due to these incompatibility, integration problems were reported, and still are, and is used as one of the main arguments used to promote Ethernet on the field level. Moreover, the

spread of Ethernet on the lower levels is also supported by the paradigm of interoperability between the enterprise systems and the control systems to interconnect them on a same network [Sauter 2010].

Nowadays, a large number of ICSs operate with Ethernet networks on the lower levels. In addition to this, most of the devices in the control system are now based on a digital architecture. From now on, sensors and actuators are subject to a digitization to improve the amount, the quality of feedback information, and on a side note the flexibility. Furthermore, the paradigm of Industry 4.0 follows this trend by introducing additional smart field devices [Alcácer & Cruz-Machado 2019, Liu *et al.* 2019, Gungor & Hancke 2009, Vitturi *et al.* 2019].

To respond to the productivity problem, the responsiveness of ICS has been optimized by improving the flow of information between the field devices and the control system. This evolution has conducted the ICSs to have mostly a digital architecture closed to the Information Technology (IT) systems on the communication and the devices, i.e. digital entities. However, by introducing digital entities operating on standard architecture mostly IT-based inside the process control levels of the ICS, a new root cause of failure has appeared: the Malevolence.

1.4 Malevolence: an intentional root cause

In this section, the malevolence is presented as a new root cause, that can lead to a process failure. This root cause is driven by an attacker with several objectives, and among them the violation of the process integrity. Specific characteristics of this root cause are then presented. First and foremost, attacks against ICS are briefly explained, and the attacks objective of concern in this thesis is focused on.

1.4.1 Violation of the process integrity

Different attacks against ICSs have been reported [Hemsley & Fisher 2018a], and are classified in two groups as follows [Gisel & Olejnik 2018]:

- **Information Operations** encompass the activity of reconnaissance, surveillance, and exfiltration of data and information about the ICS. It includes (i) stealing and disclosing sensitive information, (ii) cyber frauding (e.g. illegitimate financial transfer), and (iii) collecting information about the ICS operations (e.g. espionnage) [OECD 2017] in order to launch Effects Operations.

- **Effects Operations** encompass the activity of generating effects on the internal structure of the ICS. It includes (i) tampering with data integrity (e.g. deletion, modification), (ii) affecting availability of the devices (e.g. disabling their operations for short or prolonged periods of time), and (iii) causing physical effects (e.g. damaging the process leading potentially to human loss and environment damages). The Effect Operations are the attacks that violate the process integrity, as defined below.

In this thesis, only the attacks violating the process integrity of the ICS are of concern [Escudero *et al.* 2018]. As the internal structure controls a process as described in Section 1.2, attacks generating effects on the internal structure affect inherently the process control. Therefore, the process is controlled in a different way than the one imposed by the control system. As shown in past ICS attacks [Sicard *et al.* 2018], these attacks aim at damaging the process by violating the process integrity. In the context of ICS attacks, we define the process integrity violation as follows.

Definition 1. *The **process integrity violation** is the intended alteration or destruction of the physical system (e.g. robot, lathe, milling machine) through its control, such as for example the intended transgression of physical constraints (e.g. space shared in mutual exclusion between two robot arms, speed limit of an electric motor).*

Violation of the process integrity leads to a fault in the process. Unlike the root causes of failures studied in the literature, and seen in Section 1.3, the root cause of this failure is intended. We propose to call this new root cause: Malevolence.

1.4.2 Malevolence

1.4.2.1 Malicious act

As shown in past ICS attacks [Sicard *et al.* 2018, Hemsley & Fisher 2018a], one of the major threats the ICSs are facing is the capability for an attacker to penetrate and manipulate the control system to violate the process integrity. This malicious act is proposed to be defined as follows.

Definition 2. *A **malicious act**, commonly called an attack, is a set of intended actions executed on the control system of an ICS. These intended actions are launched by an organization, i.e. the attacker, with the primary objective of violating the process integrity to cause a fault or a failure in the process. This fault can be classified into different degrees of severity:*

- ***Degradation of the delivered service (fault)**: the process in whole or in part partially delivers the desired service, i.e. the delivered service deviates from the desired service, (e.g. constant bias or transient overshoot of desired angular velocities [de Sá et al. 2017]).*
- ***Temporarily inability to deliver the service (failure)**: the process in whole or in part temporarily loses its ability to deliver the service (e.g. intermittent failure of communication with the equipment [Ylmaz et al. 2018, Long et al. 2005, Cetinkaya et al. 2019, Amin et al. 2009]).*
- ***Permanent inability to deliver the service (failure)**: the process in whole or in part permanently loses its ability to deliver the service (e.g. permanent failure of communication with the equipment, permanent damage of the process leading to permanent interruption of operations in the ICS).*

To be realizable, the malicious act requires the penetration of the control system defined as follows.

Definition 3. *The **penetration of the control system** is the corruption of one or several digital entities of the control system.*

The corruption of a digital entity can be considered as a failure of this entity because its ability to deliver its service is not anymore guaranteed. The root cause of this failure is the malevolence. Nowadays, an attacker penetrates the control system by benefiting from the digital entities left open for maintenance and reconfiguration purposes [Luiijf 2016, Fovino 2014]. Two profiles of attacker can be pinpointed:

- **Insider threat** is "a person with legitimate access to an company's computers and networks" [Pfleeger *et al.* 2010]. In the context of ICS attacks, it is an operator with a legitimate access to the control system allowing the attacker to penetrate it.
- **Outsider threat**: the attacker penetrates the control system using various IT-based methods [Alladi *et al.* 2020] to get illegitimate access to it. Most of the time, these methods fool one or several operators to penetrate the control system (e.g. social engineering).

1.4.2.2 Characteristics of the malevolence

As the malevolence is an intended root cause of failure, it has unique characteristics. We have identified three main characteristics: the capability, the spread, and the propagation path.

The capability of the malevolence is the ability of the malicious act to successfully achieve its objective. The success of this objective clearly depends on its ability to violate the process integrity, but also to remain stealthy during the period of attack. Otherwise, the attack could be detected and isolated before it violates the process integrity, leading to a fail of the attack (e.g. Triton attack [Hemsley & Fisher 2018b]). Hence, the capability depends on the knowledge of the attacker about the ICS. This begins first and foremost with the process of knowledge, being the attack goal. In addition, the knowledge about the control system's modules is required to understand their interaction with the process. Indeed, one module can be directly interconnected with the process, or indirectly through other modules as it will be shown in Chapter 2. Moreover, the knowledge about the deployed detection systems is mandatory to guarantee the stealthiness of the attack. Clearly, the presence of detection systems limit the capability of the malevolence. Also, the quality of knowledge, i.e. incomplete, partial, complete, limit it and can even cause a fail of the attack.

The spread of the malevolence is the place in the control system where the malicious act can be performed. As the attacker targets the digital entities, the attack can hit all the modules of the control system as they are now all digitized. The malevolence is then widely spread over the whole control system. Thus, higher the number of digital entities in the control system, higher the number of potential corrupted entities are. However, higher the level of the ICS the failure occurs, longer the propagation path to violate the process integrity is. This comes from the hierarchical organization of the control system leading having upper modules than others.

The propagation path of the malevolence is the propagation path followed by the anomalies resulting from the failures of digital entities, i.e. their corruption. Unlike the other root cause of failures, the intentional nature of the malevolence conducts the anomalies to be controlled by the attacker. In case of the primary objective is the violation of the process integrity, the anomalies tend to descend in the control system to reach the process.

1.5 Conclusion

In this Chapter, the malevolence has been presented as a new-root cause of failures that aim to violate the process integrity. Unlike the other root-causes, the malevolence is intentional. From this property, we have highlighted characteristics of this intended root-cause: the capability, the spread, and the propagation path to violate the process integrity. In the next Chapter, the malevolence will be analyzed to understand how anomalies can be propagated over the control system in order to identify the digital devices that need to be secured. Obviously, this propagation intrinsically depends on the characteristics of the malevolence. Finally, we will provide a general presentation of the behavioral approaches addressing the malevolence.

Chapter 2

Analysis of the malevolence and behavioral approach

Contents

2.1 Introduction

This Chapter analyzes the malevolence with the aim to understand how anomalies can be propagated over the control system to potentially violate the process integrity (Section 2.3). This Chapter shows that the controllers need to be hit by the malicious act to violate the process integrity. This result is then shown on the Stuxnet attack (Section 2.4). Finally, a general presentation of the behavioral approaches addressing the malevolence is presented (Section 2.5). First and foremost, the ICS functioning is detailed based on the main digital devices (Section 2.2).

2.2 Preliminary: detailed ICS functioning and penetration point

2.2.1 Functions of Control

Each digital entity of the control system processes several Functions of Control to achieve their control mission. This functional view is inspired from the function blocks description of distributed control systems [Dai & Vyatkin 2012, M-SYSTEM 1990]. We define a Function of Control as follows.

Definition 4. *A **Function of Control (FC)** is a high-level abstraction procedure processed by a digital entity to achieve its control mission.*

For instance, a PLC achieves its control mission by processing the following FCs [Bolton 2015]: **reception** of status feedbacks from sensors, **computation** of decisions according to the PLC's control program, and **transmission** of the decisions to actuators.

In the following we study the FCs of both digital devices and digital communications, and their functioning. In addition, their digital architecture is considered to show the penetration points.

2.2.1.1 Digital devices

A digital device operates on a digital architecture in which FCs are processed to achieve its control mission. The digital architecture comprises three layers: a **Hardware (Hw.)**, a **Firmware (Fw.)**, and an **Application (App.)**. These layers are further detailed in [Mano 1992]. The digital architectures are either (i) computer-based which is a standard IT architecture including a dedicated App. layer, or (ii) embedded-based with only dedicated layers for the process control. In this thesis, only the layers dedicated to the process control are considered, so the Hw. and Fw. layers on computer-based systems are not considered in this thesis.

The digital devices are the core elements of an ICS. They achieve their control mission by interacting with each other. Each digital device reacts to input decisions from its supervisors, transmits output decisions to its subordinates, monitors their execution, and transmits status feedback to its supervisors [Jones & McLean 1986]. Each digital device processes a set of FCs, belonging to the following list inspired from the functions of distributed computer control systems [Steusloff 1984, Syrbe 1978, Kramer *et al.* 1984] :

- **rec** is the function of receiving and delivering the content of messages coming from other devices.
- **comp** is the function of processing the content, and computing output decisions.
- **tran** is the function of formatting and transmitting messages to other devices.
- **stor** is the function of storing the variables representing the behavior of the process and/or the devices.
- **disp** is the function of displaying valuable information to the operators and getting their decisions, both about the behavior of the process and/or the devices.

2.2.1.2 Digital communications

The digital communications denoted Communication (Com.) gather multiple devices (e.g. gateway, switch, router) which are opaque from the process control: the function of exchanging messages matters. Hence, the communications are a mean for exchanging messages, i.e. decision and status feedback, between devices.

The messages can be broken down into (i) **ctrl** for controlling the process (work order, setpoint, command, request, instruction, sensor reading, status, response), and (ii) **cnf** for configuring the digital devices, i.e. architecture reconfiguration, change of status. Both (ctrl) and (cnf) are the FCs of the communications.

2.2.1.3 Penetration points

As explained in Section 1.4, the malicious act requires the penetration of the control system. Analysis of ICS attacks shows that this penetration consists in corrupting the architecture of a digital entity, i.e. a digital device or a digital communication, to gain access to the FCs. Note that for digital communications the corruption of the architecture of their devices are not studied as they are not concerned in this thesis. Additional considerations will be stated later.

2.2.2 Main digital devices

To study how anomalies can be propagated over the control system, it is required to describe in depth the ICS functioning based on the main digital devices in charge of controlling the process. In the following, notice that the control coordinator (level 2) is merged with the supervisory control (level 3). Therefore, the ICS architecture in Figure 1.2 can be restated in Figure 2.1 with the main digital devices described as follows.

2.2.2.1 Local control (level 1)

It includes cells of controllers containing Continuous-Variable Dynamical System-based controller (CVDS controller) (e.g. Proportional-Integral-Derivative controller (PID controller)), Discrete Event Dynamical System-based controller (DEDS controller) (e.g. PLC), and HMI attached to them.

• **CVDS controller (CVDS)** is a continuous-variable controller receiving setpoints (**rec**) from a DEDS controller and/or an HMI to adjust the process behavior. It computes continuous-variable commands from its continuous-variable control law (**comp**), it transmits them to actuators (**tran**), it monitors their execution by receiving sensor readings from sensors (**rec**), and it stores the sensor readings with its status in a shared memory (**stor**). For the monitoring purpose, it can receive requests (**rec**) from a DEDS controller and/or an HMI to have access to its shared memory. In response, it reacts by transmitting the requested variables (**tran**) or by making them available in its shared memory (**stor**). Usually, an estimator is also part of (**comp**) in order to overcome the lack of sensors for describing the process behavior [Hoo *et al.* 2003, Pivarčiová & Qazizada 2018, Štefan Kozák 2014].
The CVDS controller is an embedded-based system with a reprogrammable App., i.e. control program, including a control law (comp) and a shared memory management (stor). However, the Fw. is generally designed only once by the manufacturer itself. Hence, we assume that the Fw. cannot be corrupted.

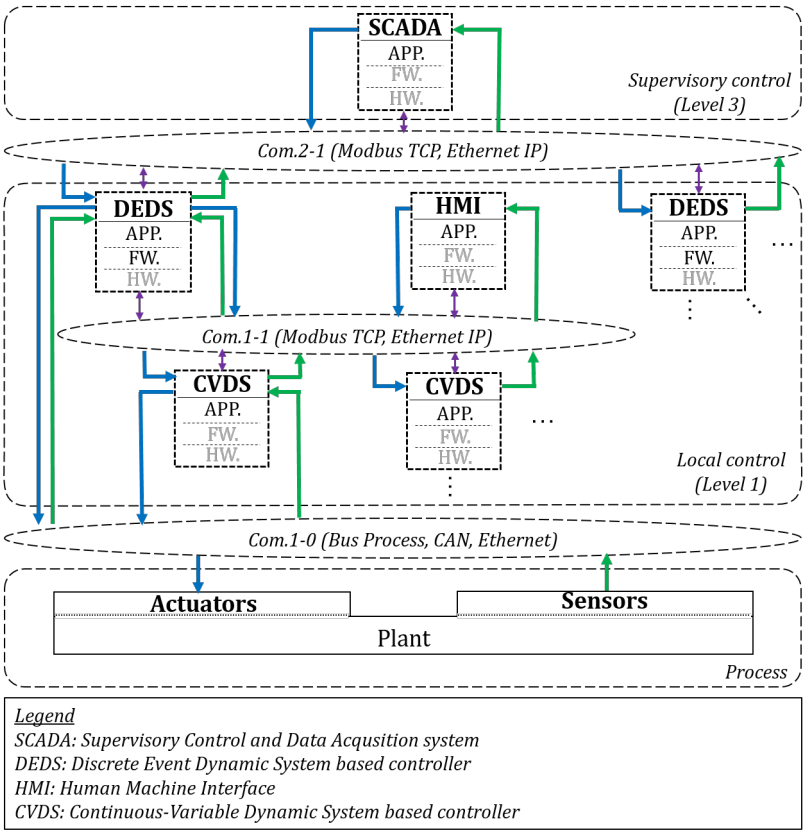


Figure 2.1: Reduced ICS architecture - Main digital devices (bold), their vulnerable architecture layers (Hw., Fw., App.) studied (in black) and not studied (in grey), and their exchanged messages: decision (→), status feedback (→), and monitoring (↔)

• **DEDS controller (DEDS)** is a discrete-variable controller receiving work orders (**rec**) from the Supervisory Control And Data Acquisition (SCADA) and/or HMI to adjust the process behavior. In reaction, it computes discrete-variable commands, i.e. command or setpoint, from its discrete-variable control law (**comp**), it transmits them to whether actuators or CVDS controllers (**tran**), it monitors their execution by receiving whether sensor readings or CVDS controllers' variables (**rec**), and it stores them with its status in a shared memory (**stor**). For the monitoring purpose, it can receive requests (**rec**) from a SCADA and/or an HMI to have access to its shared memory. In response, it reacts by transmitting the requested variables (**tran**) or by making them available in its shared memory (**stor**) [Bolton 2015].
The DEDS controller is an embedded-based system with a reprogrammable App., i.e. control program, including a control law (comp) and a shared memory management (stor); and a reprogrammable Fw. being the interface between the App. and the Hw. (rec, comp, tran, stor) [Pavesi *et al.* 2019, Bradbury 2012, Gonzalez *et al.* 2019, Valentine & Farkas 2011]. In some cases, the DEDS controller is a computer-based system with a reprogrammable App. embedding the control law [Milinković & Lazić 2012].

• **HMI (HMI)** receives decisions (**disp**) from operators to adjust the process behavior or the controllers' status. In reaction, it computes the appropriate decisions, i.e. work order or setpoint, (**comp**), and it transmits them to the appropriate controllers (**tran**). Moreover, the HMI continuously monitors the attached controllers by transmitting them requests (**tran**), receiving their responses (**rec**), and storing the controllers' variables in a database (**stor**). Then, the HMI computes (**comp**) valuable information about the local behavior of the process and of the local control, and displays it (**disp**) to the operators [Zolotová & Landryová 2000, Tendjaoui *et al.* 1991, Wucherer 2001].
The HMI is a computer-based system with a reprogrammable App. which is a dedicated program (comp, stor, disp). Often based on windows platforms, this type of device is considered highly vulnerable [Mcgrew 2013, McGrew & Vaughn 2009, Morris *et al.* 2013, Gonzalez *et al.* 2019]. As it is a computer-based system, the Hw. and Fw. are not considered in this manuscript.

2.2.2.2 Supervisory control (level 3)

• **SCADA system** receives instructions (**rec**) from operators or higher layers of ISA95 (not studied in this manuscript) to adjust the process behavior or the controllers' status. In reaction, it computes the appropriate work orders (**comp**), and transmits them to the appropriate controllers (**tran**). Moreover, the SCADA continuously monitors the attached controllers by transmitting them requests (**tran**), receiving their responses (**rec**), and storing the variables in a database (**stor**). Then, the SCADA computes (**comp**) valuable information for the operators about the global behavior of the process and the local control, and displays it (**disp**) [Daneels & Salter 1999]. Some ICSs where the controlled plant is a single entity (e.g. power grid) present specificities: (i) the control of the process is mainly deported to the SCADA where the DEDS controllers serve mostly as remote stations for transmitting the sensor readings to the SCADA, and thus (ii) an estimator is equipped in the SCADA (**comp**). The contributions depending on these specificities will be mentioned with the symbol ⋄ after the citation [Prada 2013, Juliani Correa de Godoy & Garcia 2017, Zolotová & Landryová 2000].
The SCADA is a computer-based system with a reprogrammable App. (comp, stor, disp) [Kim 2012, Upadhyay *et al.* 2020, Shapiro *et al.* 2011, Cagalaban *et al.* 2010, Fovino 2014, Gonzalez *et al.* 2019, Samtani *et al.* 2016]. Often based on windows platforms, and fully connected to the enterprise network, this type of device is considered highly vulnerable (Fw. and Hw.). As it is a computer-based system, the Hw. and Fw. are not considered in this manuscript.

2.2.3 External devices

Other devices are involved in the design or the monitoring of the ICS. They are external devices as they are not in charge of the process control.

2.2.3.1 Engineering Workstations (EWs)

Engineering Workstation (EW) are computer-based systems interconnected intermittently with the ICS to reconfigure the digital devices through configuration messages

exchanged over a communication, often the Com.2-1.

2.2.3.2 Anomaly detectors

Most of the CVDS and the DEDS controllers, and the SCADA ($\diamond$) are configured with an anomaly detector to detect faults in the process [Dorr *et al.* 1997, Combacau *et al.* 2000]. The basic idea is to forecast the process behavior with an estimator. Then, hypothesis testings based on characteristics of the residual, i.e. difference between the sensor readings and the estimation, are checked to establish either the presence of an anomaly or not.

2.3 Analysis of the malevolence

After having detailed the ICS functioning from a functional view of the main digital devices, the malevolence is analyzed. Owing to its intentional nature, anomalies are controlled by the attacker to successfully achieve its objective of violating the process integrity. To understand how anomalies can be propagated over the control system to reach the attack objective, we propose a methodology. It is based on the FCs the digital entities are processing, to ensure their control mission in the ICS.

2.3.1 Control-based attack model

From the past attacks analysis, we have established a control-based attack model. This attack model describes how the anomaly propagates once a device architecture or a communication is corrupted. This attack model is based on the FC presented in Section 2.2.

2.3.1.1 Definitions

The four stages of the malicious act are defined as follows.

- **(I) Corruption** is the modification of a device architecture (Hw., Fw., App.) or a communication, after having gained access to it. The corruption allows the manipulation of one or several FCs by the attacker. This stage leads to a failure of the digital entity as it is not anymore able to achieve its control mission, i.e. its service.

In this analysis, we do not consider the Hw. corruption as it requires to have a physical access to the ICS, which one has already a control access policy restraining such corruption; we consider also that all the devices do not contain hidden functions at the time of their manufacture.

- **(II) Manipulation** is the modification of the normal processing of a corrupted device's FC by an attacker. The Manipulation of a FC results in an anomaly, i.e. a fault.

- **(III) Propagation** is the transmission of an anomaly from a device's FC to another device's FC through the ICS architecture.

- **(IV) Misbehavior** is the anomalous behavior of a device's FC resulting from the propagation of an anomaly to this FC input. The Misbehavior of a FC results in a new anomaly in its output.

2.3.1.2 Attack model explanations

The four stages of an attack take place as follows. Once an entity of the control system is corrupted (**I**), the attacker can manipulate some of their FCs. A manipulation modifies the normal processing of a FC, and it produces anomalies (**II**). The produced anomalies are then propagated to other devices through the exchanged messages (**III**). The devices, that are exposed to these anomalies, misbehave from the propagation (**IV**). The normal processing of their FCs is not modified, but the received anomalies do produce new anomalies in the outcome of each exposed FC, as they are processing anomalous input messages. Then, these new anomalies propagate further following the same stages. The four stages are preceded and initiated by Information Operations attack to bring the attacker knowledge about the ICS architecture, the devices processing and the process behavior.

The stages of the malicious act have been presented. In the following, propagation paths of anomalies aiming at reach the process to violate its integrity are investigated.

2.3.2 Process integrity violation

In Section 2.2 the interactions between the devices have been modeled by considering the ICS is arranged following the ISA95 architecture. The attacks propagation map proposed in Table 2.2 is built from (i) the vulnerable architectures of devices and their processing, to define the structure of the table (column and row labels); and (ii) the attack model combined with the interactions between the devices, to define the content of the Table. Indeed, the interactions between the devices are at the core of Table 2.2 and of the following analysis since it defines where the anomaly can be propagated. This table gathers:

- The vulnerable devices architectures and communications in column labels

- The FCs of the devices and of the communications in row labels

For the sake of clarity in the table explanation, we define by "item" a device architecture or a communication. To understand how to read the map in Table 2.2, the reader has to view it through the attack model presented above. The methodology is provided below and the corresponding steps are presented in Figure 2.2:

- **(I) Item corruption:** this refers to the choice of the column label defining the corrupted item. After having selected the column, the reader has to find the internal processing of the corrupted item by crossing the rows and the column sharing the same label, i.e. grey font cells (▬). Note that which architecture is corrupted matters in the analysis as multiple columns can match with the rows-column crossing: the reader has to choose the right column regarding which architecture is corrupted illustrated with a black font cell (▬) and labeled at the column label. The internal processing informs about which FCs can be manipulated if the corresponding item is corrupted.

- **(II) Item's FCs manipulation:** from the internal processing (▬), the FCs that can be manipulated are mentioned with (⋆). In the opposite, the ones that cannot be manipulated have (-). The manipulation of a FC results in an anomaly on it.

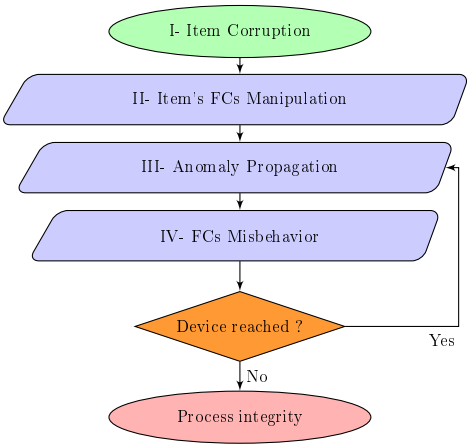


Figure 2.2: Map reading methodology

- **(III) Anomaly propagation:** once an anomaly occurs on a FC, it propagates through other FCs mentioned with ($\rightsquigarrow$). At an instant, the propagated anomaly leaves the corrupted item and continues its progression by going through one or several FCs until it reaches a device. ($\rightsquigarrow$, $\rightsquigarrow\rightsquigarrow$, $\rightsquigarrow\rightsquigarrow\rightsquigarrow$) describes the propagation path to the closest devices. A subscript has been added to the symbol to identify the propagation path when many of them are possible. Note that the anomaly propagation has to be read vertically, i.e. only the rows change but not the selected column.

- **(IV) FCs misbehavior:** the propagated anomaly reaches a non-corrupted device or the process (▬) via the exchanged messages. If this is the process, then its integrity is violated. Otherwise, if it is a device, the FCs process the messages, and then misbehave. The misbehaving FCs are mentioned with ($\blacklozenge$). Then the anomaly continues its propagation from the new anomaly resulting from the misbehaving FCs. From the current row (corresponding to the current misbehaving device), the reader has to choose the column corresponding to the device misbehaving. Note that when multiple columns exist, the reader has to choose the column with the App. device architecture. In fact, when a device misbehaves it only affects its normal processing which are defined by the App. device architecture. Once the new column is selected, the reader goes back to step **(III)** to propagate the anomalies from the misbehaving device, similarly as if the App. of the device was corrupted, i.e. with the ($\star$) representing now the misbehaving FCs creating anomalies in their output. This methodology is repeated again and again until the process is reached. If this is the case, it means that an anomaly is controlling the process, in other words the process integrity is violated-.

2.3.2.1 Case study: control law injection

In this part, we show how to use the above methodology on a case study. For simplicity purpose, a downsized version of Table 2.2 is available in Table 2.1 to emphasize the case study. This downsized version includes only the propagation paths of interest.

Consider an attacker that has corrupted the App. of a DEDS controller (e.g. PLC) by uploading a malicious control law to forge anomalous commands. Recalling that the control law corresponds to the FC: (comp) (Section 2.2).

		DEDS App.	CVDS App.
DEDS	comp	★	
	tran	⇝	
	App.	■	
Com.1-1	ctrl	⇝⇝	
CVDS	rec	⇝⇝⇝	-
	comp	♦	★
	tran	-	⇝
	stor	♦	★
Com.1-0	ctrl	-	⇝⇝
Process		-	⇝⇝⇝

Table 2.1: Reduced attacks propagation map for the case study

- **(I)** Starting from the corrupted item, the column label "DEDS App." is selected. Then, we search for its internal processing by choosing the row label of the corrupted item, and crossing them with the selected column. The internal processing is shown in (▬), and the corrupted architecture in (▬).
- **(II)** The attacker has uploaded a malicious control law that manipulates the FC (comp), which is consistent with the table since this FC can be manipulated by an attacker (⋆). An anomaly results from this manipulation: anomalous computation is processed.
- **(III)** The anomalous command propagates on the FC (tran) (⇝) to be emitted to the CVDS-controller. Then, the anomaly leaves the corrupted DEDS controller. Thereafter, the propagation continues over the Com.1-1 on the FC (ctrl) (⇝⇝) until it reaches the CVDS-controller and its FC (rec) (⇝⇝⇝). To keep simplicity in the analysis, we only analyze one propagation path over three in Table 2.2, i.e. corresponds to the one with the subscript $_{(1)}$.
- **(IV)** The anomaly transported via the message is stored in the CVDS memory (stor) and is used for the computation (comp). Hence, both FCs (stor) and (comp) misbehave (♦). Since the CVDS misbehaves, it produces a new anomaly being propagated. To study the propagation path, we choose the column label which corresponds to the misbehaving device, i.e. CVDS, with the device architecture App. From now on, the analysis is done from the CVDS App. column by going back to step (III) of Figure 2.2.
- **(III)** The anomaly goes through the FC (tran) (⇝), and is transported via Com.1-0 on FC (ctrl) (⇝⇝). As before, only one propagation path is studied which corresponds to the one with the subscript $_{(1)}$ in Table 2.2.

- **(IV)** From the propagation path, the anomaly reaches the process ($\rightsquigarrow\rightsquigarrow\rightsquigarrow$) which means the violation of its integrity.

The attacks propagation map in Table 2.2 has been proposed and a methodology to use and read it has been detailed in Figure 2.2. A case study has been studied to show how to use the methodology with a reduced attacks propagation map. As seen in this Section, the attacks propagation map is relevant to study how the process can be reached by an anomaly resulting from the corruption of a device architecture or of a communication. As a remark, only the propagation path resulting to the process integrity reachability has been detailed. However, by analyzing the other propagation paths it is possible to find out all the other devices that are misbehaving as we will see in the next part on past attacks analysis.

2.4 Application on Stuxnet attack

Apart from the case study, we describe and analyze in this part the application of the attacks propagation map in Table 2.2 on past attacks following the methodology presented above. It appears that such attacks tried to remain stealthy regarding the operators in front of the SCADA or the HMIs as an additional objective [Sicard *et al.* 2018]. In addition to the process integrity reachability, we use the methodology to analyze how the SCADA and/or the HMIs have been targeted. This is done by exploring the other propagation paths as stated in the remark above. From now on, we specify which message content part, i.e. pl or md, are involved.

2.4.0.1 Stuxnet Attack

Stuxnet attack aimed to abnormally control the centrifuges velocities, i.e. to violate the process integrity, and to hide the abnormal values of variables related to the centrifuges velocities from the operators in front of the SCADA and the HMIs. After a preliminary reconnaissance of the targeted ICS control features, the attacker has corrupted the DEDS App. by uploading a malicious control program from a corrupted EW. Such corruption allowed the manipulation of (i) the FC (comp) to forge anomalous setpoints to violate the process integrity, and (ii) the FC (stor) to store false variables representing a normal process behavior, as it should be without any attacks, in order to remain stealthy [Langner 2011].

Figure 2.3 describes the attack propagation mentioned with a reference letter (A,B,C,D) and numbers, the asterisk symbol * refers to the propagation path. It is also reported in Table 2.2 to guide the reader in the analysis. For visibility purposes, only the last propagation path's symbol is labeled in Table 2.2. Indeed, it is sufficient for the reader to read the symbol's subscript to trace the propagation path backwards. The Stuxnet attack propagation follows four distinct steps which are illustrated sequentially in the state flow (Figure 2.3). These four steps are detailed below:

1. Beforehand, the attacker has corrupted an EW (**A**), an external device (white font cell), to upload a malicious control program in the DEDS App. through the Com.2-1 (cnf-pl) (**A***).

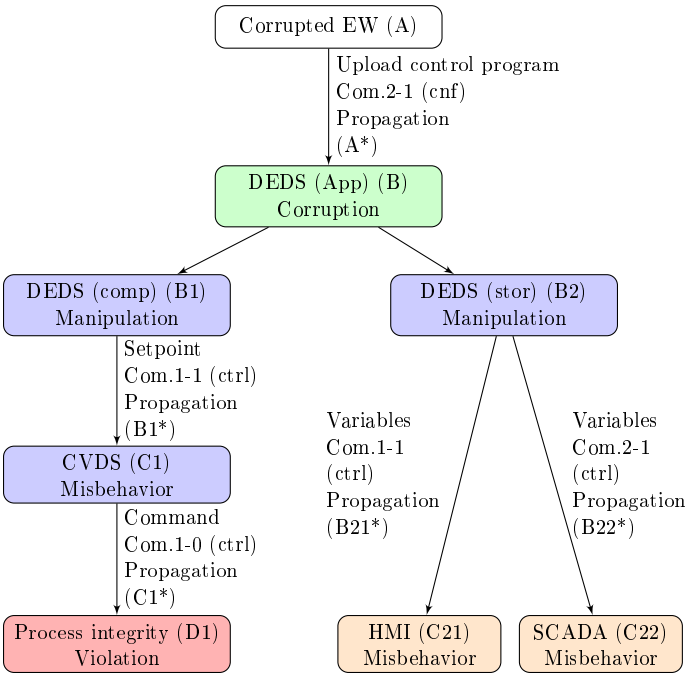


Figure 2.3: Stuxnet attack propagation

		SCADA App.	Com. 2-1		DEDS App.	B	DEDS Fw.	HMI App.	Com. 1-1		CVDS App.	C1	Com. 1-0	EW	A
SCADA	rec	-	-	⇝	⇝⇝⇝(2)	B22*	⇝⇝(2)	-	-	-	-		-	-	
	comp	★	-	♦	♦	C22	♦	-	-	-	-		-	-	
	tran	⇝	-	-	-		-	-	-	-	-		-	-	
	stor	★	-	♦	♦	C22	♦	-	-	-	-		-	-	
	disp	★	-	♦	♦	C22	♦	-	-	-	-		-	-	
	App.	■	-	-	-		-	-	-	-	-		-	-	
Com.2-1	cnf	-	★	-	-		-	-	-	-	-		-	⇝(1)	
	ctrl	⇝⇝	-	★	⇝⇝(2)		⇝(2)	-	-	-	-		-	-	
DEDS	rec	⇝⇝⇝	⇝	⇝	-		★	⇝⇝⇝	-	⇝	⇝⇝⇝(2)		⇝(1)	⇝⇝(1)	A*
	comp	♦	-	♦	★	B1	★	♦	-	♦	♦		♦	-	
	tran	-	-	-	⇝		★	-	-	-	-		-	-	
	stor	♦	-	♦	★	B2	★	♦	-	♦	♦		♦	-	
	App.	-	♦	-	■		-	-	-	-	-		-	♦	B
	Fw.	-	♦	-	-		■	-	-	-	-		-	♦	
HMI	rec	-	-	-	⇝⇝⇝(3)	B21*	⇝⇝(3)	-	-	⇝	-		-	-	
	comp	-	-	-	♦	C21	♦	★	-	♦	-		-	-	
	tran	-	-	-	-		-	⇝	-	-	-		-	-	
	stor	-	-	-	♦	C21	♦	★	-	♦	-		-	-	
	disp	-	-	-	♦	C21	♦	★	-	♦	-		-	-	
	App.	-	-	-	-		-	■	-	-	-		-	-	
Com.1-1	cnf	-	-	-	-		-	-	★	-	-		-	⇝(2)	
	ctrl	-	-	-	⇝⇝(1)(3)		⇝(1)(3)	⇝⇝	-	★	⇝⇝(2)		-	-	
CVDS	rec	-	-	-	⇝⇝⇝(1)	B1*	⇝⇝(1)	⇝⇝⇝	⇝	⇝	-		⇝	⇝⇝(2)	
	comp	-	-	-	♦	C1	♦	♦	-	♦	★	C1	♦	-	
	tran	-	-	-	-		-	-	-	-	⇝		-	-	
	stor	-	-	-	♦	C1	♦	♦	-	♦	★	C1	♦	-	
	App	-	-	-	-		-	-	♦	-	■		-	♦	
Com.1-0	ctrl	-	-	-	⇝⇝(4)		⇝(4)	-	-	-	⇝⇝(1)	C1*	★	-	
Process		-	-	-	⇝⇝⇝(4)		⇝⇝(4)	-	-	-	⇝⇝⇝(1)	D1	⇝(2)	-	

Table 2.2: Attacks propagation map for an ICS arranged following the ISA95 - From the corruption of a device architecture (■) or of a communication (▬), the potential manipulated FCs (★), the propagation path of anomalies (⇝ ⇝⇝ ⇝⇝⇝), the misbehaving FCs (♦), the process integrity (▬), the not targeted FCs and device architecture (-); the Stuxnet attack propagation (letters and numbers)

2. From now on, the DEDS is corrupted (**B**), i.e. the column and the rows change, and the attacker manipulates both (i) the DEDS (comp) (**B1**), and (ii) the DEDS (stor) (**B2**).

3. (i) The DEDS (comp) (**B1**) is manipulated:

   (a) The attacker computes an anomalous setpoint which propagates through Com.1-1 (ctrl-pl) (**B1***) to make the CVDS controller misbehave (**C1**).

   (b) From the misbehaving CVDS controller (**C1**), i.e. the column changes, anomalous commands are transmitted through Com.1-0 (ctrl-pl) (**C1***) to the process.

   (c) Hence, *the process integrity is violated* (**D1**).

4. In the meantime, (ii) the DEDS (stor) (**B2**) is manipulated:

   (a) The attacker stores false variables about the behavior of the process and of the local control in the DEDS shared memory. Then, they are propagated from the Com.1-1 (ctrl-pl) (**B21***) to the HMI (**C21**), and from the Com.2-1 (ctrl-pl) (**B22***) to the SCADA (**C22**).

   (b) Hence, the HMI and the SCADA misbehave (**C21**, **C22**) and mislead the operators about the process and local control behaviors.

To conclude this section, the analysis of the malevolence highlights the following points:

- Anomalies controlled by the malicious act follow a mandatory path to reach its objective.
- The higher the level of the corrupted module in the ICS architecture, the longer the propagation path is, i.e. various digital devices need to be crossed to reach the process.

In the past decade, the literature has proposed a large number of contributions to address the malevolence in ICS. One of the main approaches consists in modeling the normal behavior of a digital entity with respect to the evolution of the process. The normal behavior corresponds to the behavior of the entity in the malevolence free case. The next section is devoted to the behavioral approaches.

2.5 Behavioral approaches to address the malevolence

In the previous section, the malevolence has been analyzed as a mean to create an intended failure in a digital entity, i.e. the Corruption. This failure creates in term an intended fault in the behavior of the digital entity, i.e. the Manipulation. The anomalies are then propagated over the control system by misbehaving the digital devices, i.e. the Propagation and the Misbehavior.
Literature focusing on the malevolence inherits the research works from the fault-detection and the malware detection research fields, respectively for the fault created by the malevolence and the intentional characteristic of malwares. Approaches can be split into two

main groups [Vinod *et al.* 2009, Bazrafshan *et al.* 2013]: signature-based, and behavioral approaches [Sicard 2018]. In a nutshell, the signature-based approach establishes an abnormal pattern of the behavior of a digital entity representing its behavior in the malevolence case; whereas the behavioral approach models their normal behavior, i.e. in the malevolence free case. This thesis contributes to the behavioral approach.

The basic idea behind the behavioral approach is the inspection of the behavior of a digital entity, or of a digital entity with respect to the behavior of the process. Indeed, a digital entity is a module or a communication of the control system in charge of coordinating the control and controlling the process. Therefore, any decision emitted by a digital entity is a reaction of a decision from an upper module or a status feedback of a lower module about the process behavior.

Plenty of methods utilizing a behavioral approach to tackle the malevolence have emerged in the literature. Inspired from the classification of fault-detection methods [Venkatasubramanian *et al.* 2003, Nguyen 2015], we propose to split these methods into three categories:

- **Knowledge-based methods** utilizes explicit knowledge about the behavior of a digital entity, or of a digital entity with respect to the behavior of the process. This knowledge can come from technical specifications of the control system design, from functional and structural analysis of the control system, or from the historical data. For instance, [Lin *et al.* 2013] [ctrl] models specifications of the DNP3 protocol that might be violated by the malevolence. Besides, [McLaughlin 2013] [DEDS-comp] models the PLC's control program to detect inconsistency in the computed decisions if it is subject to malevolence. In [Carcano *et al.* 2009, Fovino *et al.* 2010] [DEDS-tran] the process behavior is modeled with respect to the PLC behavior, to reveal inconsistency.

- **Data-driven-based methods** utilizes learning techniques to associate a set of decisions and/or status feedback to the behavior of a digital entity. These techniques use data collected from the current ICS operations, or from historical data. For instance, [Linda *et al.* 2009, Kalech 2019] [ctrl] model the normal behavior of communication exchanges including source-destination, size and periodicity of exchanged packets, and their probability of occurrence from network properties (e.g. response time and loss of packets) in [Ponomarev & Atkison 2016] [ctrl]. Besides, [Hadziosmanovic *et al.* 2013] [DEDS-stor] models the normal range of the variables stored in the PLC. Again on the PLC, [Rrushi & Kang 2009] [DEDS-stor] models its interaction with the process to describe the evolution of the stored variables.

- **Behavior-based methods** utilizes physical knowledge about the process behavior to model it, or to model its interaction with other digital entities. In addition to this, it can utilize information knowledge about digital entities to model it. For instance, [Goldenberg & Wool 2013, Kleinmann & Wool 2017, Hahn & Govindarasu 2013] [ctrl] model the communication exchanges between an HMI and a PLC including the size and periodicity of exchanged packets, and the data transported. In [Erez & Wool 2015] the behavior of the variables stored in the PLC is modeled.

From the model of the process behavior, [Hu *et al.* 2018b, Marelli *et al.* 2018] establish properties of the status feedback characterizing the malevolence (e.g. non-randomness, distribution properties). [Koutsandria *et al.* 2014, Garcia *et al.* 2016, Caselli *et al.* 2015, Sicard *et al.* 2019] models the behavior of the process with respect to the PLC behavior to detect if the transmitted decisions will conduct to a violation of the process integrity. In addition, [Beaudet *et al.* 2020] proposes an implementation on a testbed for detecting anomaly in the control signal transmitted by a DEDS-controller [ctrl] Similarly with the CVDS-controller, [Sridhar & Govindarasu 2014] models the behavior of the process with respect to the controller behavior to detect malevolence based on the residual on the sensor readings [ctrl]. Differently, [Koucham *et al.* 2018] extracts normal behavior features of the communication using a data mining method. It obtains temporal safety properties that must be satisfied in the attack-free case [ctrl-pl]. With a similar model, [Mo & Sinopoli 2016a, Murguia *et al.* 2017, Murguia *et al.* 2020a] investigate the impact of the malevolence on the process integrity [ctrl]. [Yaseen & Bayart 2017] proposes to monitor the CVDS-controller behavior from its input and its output. It allows the detection of a control signal manipulation from a compromised controller. Lastly, [Thuillier *et al.* 2020] uses an Interval Finite Memory Observer allowing the estimation of states on an uncertain process over a temporal window. It provides an upper and lower bounds of estimate values due to the process uncertainties. This work aims at detecting and correcting anomalous sensor readings. The detection occurs when the sensor readings are outside of the estimation bounds [CVDS-rec].

2.6 Conclusion

This Chapter has shown that multiple paths exist for the anomalies to reach the process integrity. In addition, whatever the path taken, the anomalies must cross the controllers (DEDS or CVDS) to reach it. This analysis has then been shown on a real case, the Stuxnet attack. Although this analysis shows the reachability, it does not provide formal methods to prove it, neither to prevent anomalies to reach the process integrity. This thesis takes place within this framework. As we will see later, we provide formal methods for proving the reachability of the process integrity for a type of attacks. These methods are then extended to find restrictions on the control in order to guarantee the non-reachability of the process integrity. The methods allow the analysis and the prevention of malevolence.

Chapter 3

Research problem

Contents

3.1 Introduction

As we have seen in Chapter 2, there exist multiple paths for the anomaly to reach the process integrity. However, all the origin of these paths differ. It can come from the corruption of the interaction with the controllers, i.e. SCADA, HMI, Com.2-1, Com.1-1, or from the corruption of the controllers itself and their respective communication with the process, i.e. Com.1-1. This thesis addresses the malevolence issue from the controllers, in particular the CVDS-controllers.

This Chapter is devoted to the research problem this thesis contributes. Section 3.2 describes the control system we consider in this thesis and the mathematical model utilized for describing the process behavior. Thereafter, Section 3.3 reviews briefly the literature to outline the main trends in analysis and prevention of anomalies. Then, Section 3.4 and Section 3.5 explains why our research problem is focused on anomalies in the control signal. Section 3.6 presents our problems statement, and Section 3.7 presents the methods we derive to address our problem statement and the differences with the literature. Lastly, Section 3.8 highlights the scientific issues this thesis addresses.

3.2 Process control with CVDS-controller

3.2.1 Description of the system in our focus of interest

In this thesis, we consider a control system including a CVDS controller (C) that controls a Process (P) via a transmission (tran) and reception (rec) channels. This control system is coordinated by the upper modules of the ICS: it receives setpoints being the industrial objectives. The controller imposes the desired service, often called the desired system response, by transmitting a control signal $u(t)$ to the process, and receives a partial or complete measure of the state $x(t)$ to update the control input. In addition to this, a typical anomaly detector receives the measured state $x(t)$ and the transmitted input $u(t)$ to detect a service degradation that could occur from a fault in the process. Theory of automatic control [Dorf & Bishop 2000] commonly represents this control system with a block diagram representation in Figure 3.1.

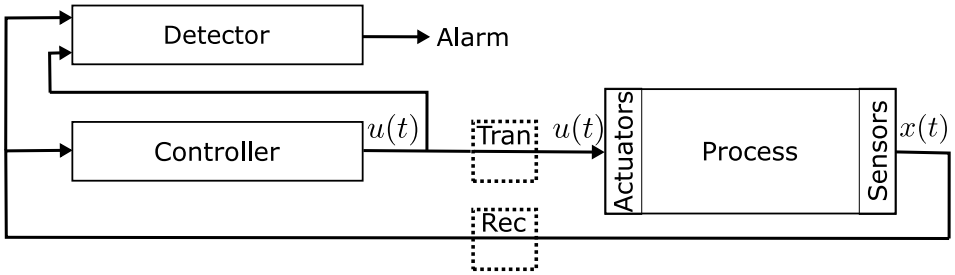


Figure 3.1: Considered control system

In this representation, the digital entities, that can be subject to malevolence, are: the controller, and the transmission and reception channels. However, the anomaly detector is assumed as not attackable. The process to be controlled by the controller is considered to be a dynamical systems. A dynamical system is "a particle or ensemble of particles whose state varies over time and thus obeys differential equations involving time derivatives" [Nature].

3.2.2 Behavior of dynamical systems

In order to provide methods to tackle the issue of malevolence in the control of dynamical systems, this thesis focuses on quantitative mathematical models of the behavior of dynamical system. Dynamical systems are represented mathematically by a set of simultaneous differential equations given in (3.1). It describes quantitatively the relationships between the system variables. They are obtained by using the physical laws governing the process (e.g. mechanical, electrical, fluid, thermodynamic).

$$\dot{x}(t) = f(x, u, t) \tag{3.1}$$

where $t \in \mathbb{R}$ is the time, $x(t) \in X \subset \mathbb{R}^n$ is the state, $u \in U \subset \mathbb{R}^m$ is the control input, and f is a function. Notice that $\mathbb{R}$ is the set of real numbers, $\mathbb{R}^n$ is the set of real n vector, and $\mathbb{R}^{n \times m}$ is the set of real $n \times m$ matrices. This notation will be recall later in Chapter 4.

In this thesis, we restrain our research works to dynamical systems modeled as continuous-time invariant affine system given in (3.2).

$$\dot{x}(t) = Ax(t) + Bu(t) + a \tag{3.2}$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^m$ is the input, $A \in \mathbb{R}^{n \times n}$ is the state matrix, $B \in \mathbb{R}^{n \times m}$ is the input matrix, and $a \in \mathbb{R}^n$ is the affine term.

An affine systems is a linear system with an affine term. In fact, the affine term can be seen as a constant control input acting on the dynamical system. Thus, one can extend the input $u(t)$ to simplify the model in (3.2). Instead of this, the model is proposed to be simplified by extending the state in order to fit with the formalism of the class of Piecewise affine (PWA) systems [Hassibi & Boyd 1998] for future research works. By defining the extended state $\tilde{x}(t) = [x^\top(t), 1]^\top$, the model in (3.2) can be reformulated as:

$$\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) + \tilde{B}u(t) \tag{3.3}$$

with

$$\tilde{A} = \begin{bmatrix} A & a \\ 0_{1,n} & 0 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} B \\ 0_{1,m} \end{bmatrix}. \tag{3.4}$$

In this section, we have described our focus of interest: the malevolence in the control of dynamical systems. From this control view, we are interested in the signals: the measurement signals, and the control signals. In the next section, we review the literature around the malevolence in the control of dynamical systems.

3.3 Brief literature review of malevolence in dynamical system

This section gives a brief literature review to outline the main trends in the literature of malevolence in dynamical system with a state-space representation. Clearly, the objective is not to provide an exhaustive list of methods, but show what is investigated in the literature.

The literature contributions can be classified into three groups as follows.

- **Vulnerability assessment** contribution explores control-based security weak points that might be abused by the malicious act (e.g. impact of manipulation on the process integrity, design of attacks).
- **Prevention** contribution provides design conditions to limit or avoid the malicious act feasibility (e.g. minimum set of sensor to secure to guarantee the state reconstruction in the attack case, configuration of anomaly detector to limit the attack impact).
- **Detection** contribution aims at revealing the presence of anomalies in the current operations of the ICS (e.g. verification of signals properties such as randomness of measurement signals, violated in the attack case).

This thesis contributes to vulnerability assessment and prevention groups. In the rest of this thesis, detection methods are thus not anymore considered.

The reviewed literature comes from the theory of automatic control. As seen previously, it considers the problem as an interconnection of two systems, i.e. the controller and the process, that reacts to signals. This systemic view leads to a lack of practical implementations of the proposed methods: for instance, measurement signals can be investigated from the sensors, the reception channel [Com.1-0-ctrl], or the controller [rec, stor, comp], however the model does not explicitly consider it. Therefore, the methods are classified either on the measurement signals or the control signals.

3.3.1 Methods based on the measurement signals

Vulnerability assessment

Three main trends appear in the literature: (i) quantification of the impact of sensor reading manipulation on the process state, (ii) design of stealthy sensor reading injection attack, and (iii) design of stealthy sensor reading Denial of Service (DoS) attack.

(i) [Bai *et al.* 2017a] considers the manipulation of a single sensor reading. From an information-theoretic approach, a graded notion of attack stealthiness is provided. Also, the fundamental limitations for the detection with an anomaly detector are given. Then, an optimal attack is designed to maximize the performance degradation for a given stealthiness level. Finally, [Murguia *et al.* 2020b] proposes two security metrics to quantify the impact on the performance degradation, based on the reachable set of the process state.

(ii) [Guo *et al.* 2018] considers the stealthy manipulation of a single sensor reading as an affine function of its legitimate value. An optimal attack is designed to maximize the performance degradation and to achieve the desired level of stealthiness expressed from the Kullback-Leibler divergence (K-L divergence) on the residue. Similarly, a more general attack is proposed in [Li & Yang 2019]. Their attack achieves the largest performance degradation of the process control. In addition, [Guo *et al.* 2019] considers a similar attack but with an extra sensor reading for measuring the process behavior. The optimal attack is designed to evade a static Chi-squared anomaly detector. Besides, [Cheng *et al.* 2019] designs event-based stealthy sensor readings manipulations to maximize the performance degradation under a fixed average communication rate to remain stealthy.

(iii) [Zhang *et al.* 2016] designs an optimal DoS attack subject to energy-constraint to maximize the performance degradation of the process control. The proposal considers the manipulation of a single sensor reading during a limited period of time. In fact, a fixed energy amount is consumed for each manipulation. Also, the impact on the control loop stability is investigated. Under similar considerations, [Zhang *et al.* 2018] designs an optimal DoS attack taking into account the energy allocation. In particular, a real-time algorithm is provided to dynamically adjust the attack energy.

Prevention

Two main trends appear in the literature: (i) conditions to guarantee the integrity of the state estimator, and (ii) configuring anomaly detectors to mitigate stealthy sensor reading manipulation.

(i) [Fawzi *et al.* 2014] characterizes the fundamental limitations of an estimator to correct invariant sparse manipulated sensor readings. In particular, it is shown that the state reconstruction is not possible if more than half the sensor readings are manipulated. An algorithm inspired from techniques in compressed sensing is provided to correct the manipulated sensor readings, when the number of manipulations is smaller than a threshold. In addition, [Chang *et al.* 2018] considers a variant set of sparse manipulated sensor readings. A secure estimator is proposed and the maximum number of manipulated sensor readings that can be corrected by the estimator is proved. In another way, [Hu *et al.* 2018a] proposes a security criterion to guarantee the security of the estimator, i.e. estimation error is bounded, when the number of manipulation is smaller than a threshold. Then, the set of normal sensor readings to guarantee the state reconstruction is provided.

(ii) [Umsonst *et al.* 2017] considers time-limited manipulations on the sensor readings and the commands. The optimal stealthy attack to maximize the performance degradation is designed for both stateless and stateful anomaly detectors. It is shown that the stateful one better mitigates the stealthy attack [rec]. Furthermore, [Umsonst & Sandberg 2018] compares the performance degradation induced by sensor readings manipulations for stateless (Chi-squared) and stateful (Cumulative Sum (CUSUM), Multivariate Exponentially Weighted Moving Average Filter (MEWMA)) anomaly detectors. It is pointed out that the MEWMA anomaly detector better mitigates the attack than the Chi-squared anomaly detector. Differently, [Murguia & Ruths 2019] proposes a systematic procedure configuring a CUSUM anomaly detector based on statistical properties of the residual. The sufficient conditions on the bias parameter are provided to guarantee the CUSUM sequence being bounded. [Miloševič *et al.* 2017] proposes a criterion to select the best combination of sensors to be secured in order to mitigate sensor readings manipulations. The defender's objective is to guarantee an upper bound of the performance degradation. Besides, in the power grid domain, [Zhang *et al.* 2019a] investigates the impact of sensor readings manipulations on the microgrid system stability. A stable region is defined and the sufficient conditions to guarantee the system stability are given.

3.3.2 Methods based on the control signals

Vulnerability assessment

Three main trends appear in the literature: (i) quantification of the impact of control signal manipulation on the process state, (ii) design of stealthy control signal injection attack, and (iii) design of stealthy sensor reading DoS attack.

(i) [Mo & Sinopoli 2016b] computes the inner and the outer-approximations of the reachable set of the process states by the means of invariant ellipsoids. Note that it also considers sensor reading manipulation. In addition, [Bai *et al.* 2017b] describes the ability of an attacker to modify a single command; and provides conditions to have the optimal performance degradation. [Milošević *et al.* 2019] quantifies the impact of

stealthy sensor readings/commands manipulations upon critical states. Two metrics are computed for describing the probability for some states to leave a safety region during a period of time. Various types of attacks are evaluated (e.g. DoS, values permutation, sign alternation) in order to find the one with the largest impact upon a set of critical states.

(ii) [Li *et al.* 2019] designs a two agents based attack applying the Least Squares Support Vector Machine (LSSVM). The CVDS controller is assumed to implement a Proportional Integrate Derivative (PID) control law. On one hand, an attack agent estimates the PID's parameters from the single sensor reading and the single command. Thereafter, the attack agent injects additional data to the command in order to deviate the process steady-state. On the other hand, a covert agent uses the estimated PID's parameters to mimic the normal steady-state by injecting additional data to the sensor reading. In addition, [Zhang *et al.* 2019b] investigates the manipulation of commands and sensor readings with a fixed additional data at each time instant. An optimal stealthy attack to maximize the performance degradation is designed.

Prevention

Only one main trend appears in the literature: (i) the restriction of the control signal.

(i) As far as we know, [Kafash *et al.* 2018] is the first and only contribution with this objective. The main idea is to quantify the impact of control signal manipulation on the process, analyze if the a dangerous zone can be reached, and restrain the control signal such that the dangerous zone cannot be reached. This work is based on the computation of the reachable set of the process. Thereafter, the same authors have extended their work in [Hadizadeh Kafash *et al.* 2018] to tackle the problem of control restriction. In fact, they have observed that by restraining the control input set, it results that the controller is too restrained to deliver the desired service.

3.4 Observation of the anomaly

As seen in the previous section, contributions focusing on the malevolence in control systems study anomalies either in the measurement signals, or in the control signals. In Chapter 2, the malevolence has been analyzed, and it has been shown that anomalies follow specific propagation paths to violate the process integrity. We propose in this section to study the issue of the anomaly observation for securing the process integrity, against a malicious act. We consider three scenarios where a single digital entity is corrupted: (i) the controller, (ii) the transmission channel, and (iii) the reception channel.

In the scenario (i) where the controller is corrupted, the attacker can manipulate the commands from the FCs (stor,comp). The abnormal command is then propagated over the transmission channel to the actuators. The process is then transformed by the actuators, so its integrity is violated. Thereafter, abnormal measurements are then sent back to the controller about the process state.

In the scenario (ii) where the transmission channel is corrupted, the attacker can manipulate the commands transmitted by the controller from the FC (ctrl). The abnormal command is then propagated to the actuators. The process is then transformed by the

actuators, so its integrity is violated. Thereafter, abnormal measurements are then sent back to the controller about the process state.

In the scenario (iii) where the reception channel is corrupted, the attacker can manipulate the measurements received by the controller from the FC (ctrl). The abnormal measurements are then propagated to the controller. The controller misbehaves: it stores and computes abnormal commands. Then, the abnormal commands are propagated over the transmission channel to the actuators. The process is then transformed by the actuators, so its integrity is violated.

We state that the study of anomalies in the measurements is too late to detect a violation of the process integrity. Indeed, the process integrity is already violated before the presence of anomalies in the measurements as seen in the scenarios (i) and (ii). Thus, abnormal measurements are the result of the malicious act on the process. In addition to this, presence of abnormal measurements do not systematically mean the violation of the process integrity. Indeed, it can also be the result of equipment failures, the most common root cause of failures in ICSs. Therefore, the root cause of failure causing abnormal measurements in the scenarios (iii) cannot be characterized. However, anomalies in the commands systematically violate the process integrity, as seen in Chapter 2. This is why it is more appropriate to study anomalies in the commands, as we can see in the scenarios (i), (ii), and (iii). However, the characterization of the root cause is not easier when anomalies are present in the commands. Indeed, it can be the result of the other root causes presented in Chapter 1.

From this observation, this thesis focuses on anomalies in the commands transmitted by the controller to the actuators. In particular, the malicious act with the primary objective of violating the process integrity to cause a permanent inability to deliver the service is focused on. To that end, we consider a malicious act that accelerates the aging of the actuators in order to cause premature failure in the process.

3.5 Aging of actuators

Aging of an actuator is a process changing the properties of the actuator with respect to the time. The type of aging of interest is the one associated with deterioration of the actuator's functional capability. Aging occurs owing to many physical stresses the actuator is subject to. They are categorized into internal and external and defined as follows [Carfagno *et al.* 1980].

- **Internal** are operational stresses meaning that they are inherent in the operation of the actuator (e.g. current, voltage, ohmic heating).
- **External** are stresses originating in the environment of the actuator (e.g. ambient temperature, vibration, humidity, oxidation) [Cox *et al.* 1995].

In this thesis, external stresses on an actuator are not considered as they cannot be controlled from the control of the actuator. Theoretical relationships between a stress and a deterioration are known for some types of materials and components. However,

for actuator mixing materials and components is often more complex. Thus the relation of degradation to stress is based on empirical knowledge, or the observation of aging as a function of time and stress magnitude.

Physical stresses are usually classified into three categories: electrical, mechanical, and thermal. They are detailed in [IAE 1989, McNeill 1983], and some internal stresses are given as follows:

- **Electrical**: electric fields, magnetic fields
- **Mechanical**: vibration, mechanical shock, i.e. bump, acceleration
- **Thermal**: temperature, temperature cycling, thermal shock, i.e. rapid change of temperature

For instance, the electric power quality referring to the quality of voltage or current is an important electrical stress (e.g. voltage and current transients, short- and long-duration voltage variation, voltage imbalance, voltage fluctuation) in power systems and electric machinery [Fuc 2008]. In the mechanical stresses, mechanical vibrations and shocks play an important role in the aging of mechanical systems [Lalanne 2013]. In particular, torsional vibration shaft due to torque harmonics has been studied in turbine-generators [Walker *et al.* 1981, Liu *et al.* 2014], and in electric machinery [Han & Palazzolo 2013, Holopainen *et al.* 2013]. In addition, mechanical overload causes an electric drive to overheat leading to a failure [Atallah *et al.* 2007]. In the thermal stresses, some applications require high peak torque, beyond the operating conditions, achieved by supplying more current to the electric drive for limited time periods. As a result, the machine is thermally overloaded and the insulation system can be degraded [Sciascera *et al.* 2016, Madonna *et al.* 2020]. Usually, aging is not only the cause of a single type of stress, but a combination of electrical, mechanical, and thermal stresses as for the degradation of the insulation system in electric machinery [Bartnikas & Morin 2004].

To conclude, this thesis is concerned about anomalies in the control signal that aim to accelerate the aging of the process's actuators [Merouane *et al.* 2020]. In the following Sections, the problem statement is presented and our research works are positioned according to the literature.

3.6 Problem statement

This thesis concerns the malevolence in the control of dynamical systems. Our interest is in malicious acts aiming at (i) accelerating the aging of a dynamical system, while (ii) delivering the desired service. (i) is the primary goal of the malicious act, i.e. violating the process integrity, and (ii) is to avoid any suspicion of fault from the anomaly detector. This is what we call a stealthy aging attack and is defined as follows:

Definition 5. *A **stealthy aging attack** drives a targeted dynamical system as in (3.3) to operate outside of the normal aging conditions while delivering the desired service.*

By accelerating the aging, a failure will occur prematurely in the dynamical system. As the malicious act focuses on the aging instead of the delivered service, the failure

could be misinterpreted and associated to a normal aging. Indeed, although reliability methods, predicting the time of failure, have been improved, they still lack in successfully predicting a failure in many cases as the aging mechanisms are complex to model and depend on various operating conditions.

As far as we know, the only contribution motivated by the stresses in a dynamical system to cause a failure is [de Sá *et al.* 2017]. It attempts to design an attack maximizing the overshoot, i.e. the peak value of a state variable during the transient response, or the constant steady-state error of the controlled variable. It considers a corrupted transmission channel and a corrupted reception channel where the malicious act manipulates the transmitted commands following an attack function. This work proposes a design method of this function, considering the attacker is aware of the controller and the process models.
Three critical remarks can be stated on this work:

Firstly, it only concerns Single-Input-Single-Output dynamical systems with an input-output relationship in the frequency-domain. The design method utilizes the root locus method for designing the attack function to make the control system underdamped, i.e. to cause overshoot. However, the work can be easily extended on the Multi-Input-Multi-Output (MIMO) system with a time-domain representation using a typical pole-placement technique from the characteristic equation of the control system.

Secondly, the design method requires the controller model. Indeed, the design of the attack function requires both the controller and the process models. That is why, the reception channel is also considered as corrupted in order to perform a system identification of both models. In a nutshell, the system identification collects the input and output signals of the system of interest to assess a fitting model, i.e. applying the same inputs, the same outputs result. But, system identification requires the structure of the model that is the unknown coefficients of the model. Both requirements on the corrupted channels and the model structure are strong.

Thirdly, the stealthiness of the attack is not considered, with respect to an anomaly detector. Indeed, if an anomaly detector is placed, the overshoot or the constant bias would be detected. By collecting the received measurement signals and the transmitted control signals from the controller side, an abnormal residual would appear, resulting probably to an alarm, i.e. depending on the hypothesis testing, [Cómbita *et al.* 2019].

This last remark is the most critical one, and clearly limits the application of their proposal. We state that this kind of attacks can be stealthy if the states related to the aging are different from the state of the service to be delivered. Therefore, we focus on the MIMO dynamical system. In addition, we assume periodicity in the production objectives as it is the case in many ICSs. We consider an attacker with the following considerations: (a) it has penetrated the control system, and corrupted the transmission channel; (b) it is aware of the process to be controlled, so it knows the process model, (c) it is aware of the periodicity in the production objectives, so that it knows the expected desired services, assuming the controller model does not change between two periodic operations. The considered control system with the attacker is shown in Figure 3.2.

In practice, such considerations are reasonable: (a) it has been shown from past ICS attacks and the literature that communication channels can be corrupted mainly from protocol vulnerabilities; (b) the process model is known by the ICS designer as

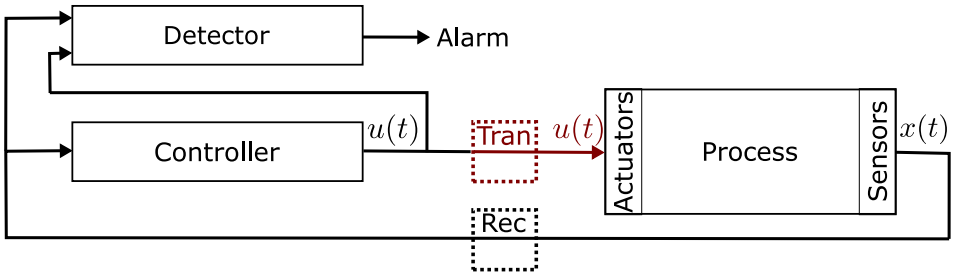


Figure 3.2: Considered control system with the corrupted channel (red)

it is required for designing the ICS and configuring the controller. As ICS applies a traceability policy, the process model should appear from the technical documentations. Thus, an attacker could gain access to it using Information Operations as preliminary attack; (c) periodicity in the production objectives depends on the variability of the ICS. Many ICSs have low variability (e.g. manufacturing production, papermill), and thus the production objectives do not change much. Hence, the attacker can perform again an Information Operations as preliminary attack to know the time evolution of the desired service. It can be known by collecting the control signals transmitted by the controller, and applying them on the process model known in (b) in order to compute the measurement signals. These measurement signals are those collected by the anomaly detector, and the state estimator of the controller, to check the hypothesis testings. This time evolution of the desired service can also be updated after having launched the stealthy aging attack. In fact, stealthy aging attacks accelerate the aging of the process by executing the attack periodically. Thus, it is reasonable to assume that in ICSs with low variability, the computational time of a stealthy aging attack is lower than the industrial objectives changes.

3.7 Scientific positioning

In this thesis, two research problems are addressed: (i) the analysis of stealthy aging attacks for quantifying the impact on the process, and (ii) the prevention of stealthy aging attacks by restraining the control signal. For (i), we propose set-theoretic methods derived from set theory to analyze the attacks feasibility. For (ii), those methods are used for restraining the control signal to make those attacks non-feasible. As we will see later, the latter has been done on a subcase of stealthy aging attacks. Thus, analysis of stealthy aging attacks (i) for the general case has been proposed with a Linear Programming (LP) method for computing the optimal abnormal input signal that maximizes the dynamical system aging.

3.7.1 Set-theoretic methods in control

Set theory is a branch of mathematics investigating sets and their properties. Many control problems can be formulated, analyzed, and solved in a set-theoretic context. Indeed, sets appear naturally when constraints, uncertainties, or design specifications are

considered. Furthermore, sets are adapted to quantify system performances (e.g. effect of persistent noise in a feedback system, state estimation error) [Blanchini & Miani 2015]. Set-theory appears in various areas of automatic control including constrained control, robust control, disturbance rejection, and robust estimation.

In this thesis, we are interested specifically in set-theoretic methods for the analysis of dynamical systems. As part of these methods, reachable set computation are the basis of many control methods including target avoidance of an adversary in a game theoretic setting [Mitchell *et al.* 2005], collision avoidance in unmanned aerial vehicles [Zhou & Baras 2015], or in fault detection for chemical reactors [Tulsyan & I. Barton 2016]. In particular, the ellipsoidal set has been largely studied in many control problems to compute reachable sets of linear systems [Blanchini 1999]. For instance, [Khlebnikov *et al.* 2011] uses invariant ellipsoids to analyze and design controller for linear time-invariant system subject to bounded exogenous disturbances.

For a dynamical system modeled as in (3.3), the reachable set is defined as follows.

Definition 6 (Reachable set, [Pecsvaradi & Narendra 1971])**.** *A (finite-time) reachable state set $\mathcal{R}(t_1, X_0)$ is the state set containing all the state trajectories of the dynamical system at a time t_1 from an initial state set X_0 with a constrained input set U.*

In the case of infinite-time, i.e. $t_1 = \infty$, the infinite-time reachable state set $\mathcal{R}(\infty, X_0)$ is the set that captures all the finite time reachable state sets from an initial state set X_0. Obviously to compute it, it is necessary for the system to be Lyapunov stable, as we will detail later; otherwise $\mathcal{R}(\infty, X_0)$ is not bounded [Gayek 1991].

In the context of the malevolence, few research works have proposed methods based on the reachable set. Our research works originate from the many works of the author encompassed in [Murguia *et al.* 2020b] where it uses the infinite-time reachable set for quantifying the impact of sensor reading or control signal manipulation on the process. Moreover, [Kafash *et al.* 2018] uses the infinite-time reachable set for restraining the control input set such that the impact of attacks avoids a dangerous set. Although this idea is not novel as it is used in safety verification and control design, it is the first contribution dealing with the malevolence. Our research work exploits this idea of restraining the control input set so that it avoids a dangerous set. However the system considered differs from the one considered in [Murguia *et al.* 2020b, Kafash *et al.* 2018]: we consider only the process without the controller, unlike the both contributions. Indeed, as seen in Chapter 2, the controller is subject to the malevolence as much as the communication channels. Therefore, we are interested in quantifying the impact of control signal manipulation by only considering the process. Thus, we search for preventing anomalies in the control signal that can come from the control system.

3.7.2 Linear optimization

Linear optimization is a method to maximize a cost function in a mathematical model whose requirements are represented by linear relationships. It is used in control problems such as path planning control to compute optimal control signals to minimize or maximize an objective function representing for instance energy consumption or actuator limitations [Puglia 2012, Yang *et al.* 2012].

In the context of the malevolence, linear optimization has not been utilized surprisingly. Thus, the closest related works are [de Sá *et al.* 2017] computing an attack transfer function maximizing the service degradation during the transient response; and [Teixeira 2019], [Zhang & Venkitasubramaniam 2017], maximizing the service degradation while being stealthy with respect to a detector. However, those contributions consider the controller in the optimization problem to escape the anomaly detector, as they degrade the service. The main difference in our work is that the attack objective is the aging of the system, instead of degrading the delivered service.

3.8 Scientific issues

In this thesis, we propose to analyze and prevent stealthy aging attacks defined in 5. The analysis consists in quantifying the impact of control signal manipulation on the process. As stated previously, we do not consider the controller model in our work, only the process model is considered.

Comparing with the literature, the first novelty in this thesis is to model the stealthiness of the attack by constraining the state to remain inside a set called the service set, as we will see in the next Chapters. In a nutshell, the state remains inside this service set when the controller controls the physical system, i.e. in the attack-free case. The state concerned with this constraint is the state that can reveal the attack. In particular, it includes: (i) the state corresponding to the service that the physical system must deliver (e.g. tracking a position reference), (ii) the state measured by the sensors, providing a feedback to the controller.

We have developed two set-theoretic methods. The first one is based on the invariant notions for computing the constrained reachable set at the infinite-time. The second one is based on convex robust simulation for computing the constrained reachable set at the finite-time. Notice that those reachable sets are referred as constrained since they are constrained by the service set. As we will see later, the second method has been developed to tackle the drawback of the invariant set-based method. In fact, as stated above, the invariant set-based method requires a condition of stability for the dynamical system of study. In our case, the system considered is the process only. But, the condition of stability is often not satisfied without considering the controller.

In addition to this, invariant-set based method loses the temporal variable in the analysis. The convex robust simulation-based method allows the presence of the temporal variable and thus, can be used to quantify at which time the dangerous set can be reached.

To summarize, the following scientific issues are addressed in this thesis:

- Modeling the attack stealthiness in set theory,
- Stability condition required from the invariant set theory,
- Loss of the temporal variable in the invariant set theory.

3.9 Conclusion

This Chapter has described the research problem addressed in this thesis. We contribute in the vulnerability assessment and the prevention of stealthy aging attacks, a new type of attack, launching on a process modeled as a dynamical system. In particular, we focus on anomalies in the control signals that reach the actuators. It can come from a corruption of the CVDS-controller or of the Com.1-1 with the process. In the next Part, our methods for analyzing and preventing stealthy aging attacks are detailed.

Part II

Analysis and prevention of stealthy aging attacks

Chapter 4

Background

Contents

4.1 Introduction

The methods for analyzing and preventing stealthy aging attacks developed in this thesis are based on dynamical system theory. This theory allows the behavior description of dynamical systems subject to input signals under some constraints. This Chapter provides the sufficient background in a compact manner to the reader for understanding the rest of this thesis.

4.2 Notations and conventions

The notations and the conventions used in the manuscript are common for the control system literature. Their description is provided in the following.

First of all, we present the notations for number sets and matrix sets. Let $\mathbb{R}$ be the set of real numbers, $\mathbb{R}^n$ be the set of real n vector, and $\mathbb{R}^{n\times m}$ be the set of real $n \times m$ matrices. Notice that $\mathbb{R}^{n\times n}$ is the set of real $n \times n$ symmetric matrices. Given a vector

$v \in \mathbb{R}^n$, $v^\intercal$ denotes its transpose. Given a matrix A, $A^\top$ indicates its transpose, $diag(A)$ is the diagonal entries of the matrix A, I_n is the identity matrix of size n, $0_{n,m}$ is the zero matrix of size n, m. The notation $A \geq 0$ ($A \leq 0$) indicates the positive (negative) semidefiniteness, i.e. all the eigenvalues of the symmetric matrix A are positive (negative) or equal to zero, whereas $A > 0$ ($A < 0$) indicates the positive (negative) definiteness, i.e. eigenvalues are strictly positive (negative). Finally, let $\mathbb{R}_m[t]$ be the set of real-valued polynomials of degree up to m in the variable t, and $\mathbb{R}_m^{n\times n}[t]$ be the set of symmetric matrix-valued polynomials of degree up to m in the variable t of size $n \times n$, with $m \in \mathbb{N}$.

The following convention, classical in automatic control, must be recalled. Generally, the signals manipulated in this manuscript are in continous-time, for example $x(t) \in \mathbb{R}^n$. Whenever this is not leading to confusions the time dependence will be dropped.

4.3 Matrix inequalities

As seen in the previous section, we focus on dynamical system described with a state-space representation using a matrix form. For this reason, it is important to recall some definitions and properties related to matrices. The following definitions and properties are highlighted.

Definition 7 (**Positive (negative) definiteness**, [Boyd *et al.* 1994]). *A symmetric matrix $P = P^\top \in \mathbb{R}^{n\times n}$ is called a negative definite matrix (respectively positive definite matrix), denoted by $P < 0$ (resp. $P > 0$), if $z^\top Pz < 0$ (resp. $z^\top Pz > 0$) for all non-zero vectors z with real entries ($z \in \mathbb{R}^n \setminus \{0_n\}$).*

Definition 8 (**Positive (negative) semidefiniteness**, [Boyd *et al.* 1994]). *A symmetric matrix $P = P^\top \in \mathbb{R}^{n\times n}$ is called a negative semidefinite matrix (respectively positive semidefinite matrix), denoted by $P \leq 0$ (resp. $P \geq 0$), if $z^\top Pz \leq 0$ (resp. $z^\top Pz \geq 0$) for all non-zero vectors z with real entries ($z \in \mathbb{R}^n \setminus \{0_n\}$)*

The methods developed in this thesis are based on the search for symmetric matrices satisfying some constraints expressed in terms of Linear Matrix Inequality (LMI). LMIs are inequalities constraining the eigenvalues of a square symmetric matrix that depends affinely (linear plus a constant term) on the decision variables.

Definition 9 (**Linear Matrix Inequality**, [Boyd *et al.* 1994]). *A Linear Matrix Inequality (LMI) is defined by the following formulation:*

$$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i > 0, \tag{4.1}$$

where $x = [x_1, x_2, \cdots, x_m]^\top \in \mathbb{R}^m$ is the vector of decision variables, and $F_i \in \mathbb{R}^{n\times n}$, $i = 0, \cdots, n$ are given symmetric matrices, i.e. $F_i = F_i^\top$. The components x_i, $i = 0, \cdots, m$, of this vector are called decision variables. The inequality means that $F(x)$ is a positive definite matrix, which can be rewritten as follows:

$$\forall u \in \mathbb{R}^n, \quad u^\top F(x) u > 0, \quad u \neq 0 \tag{4.2}$$

Remark 1. *The LMI in (4.1) is a convex constraint on x, i.e. the set $\{x|F(x) > 0\}$ is convex.*

In most control applications, LMIs are functions of matrix variables rather than scalar valued decision variables. This means that we consider inequalities of the form (4.1) where the decision variables becomes $X \in \mathbb{R}^{m_1 \times m_2}$ [Boyd *et al.* 1994]. As an example, consider the well known Lyapunov inequality $F(X) = A^\top X + XA < 0$ where $A \in \mathbb{R}^{n \times n}$ is assumed to be given, and $X \in \mathbb{R}^{n \times n}$ is a symmetric matrix variable. The phrase "the LMI $F(X) < 0$ in X" means that the matrix X is a decision variable. Note that the LMI can be rewritten in the form of (4.1) as follows. Let $X_1, \cdots, P_m$ be a basis for symmetric $n \times n$ matrices ($m = n(n+1)/2$). Then, take $F_0 = 0$, and $F_i = -A^\top P_i - P_i A$.

The form of an LMI is very general and many inequalities can be expressed as LMIs such that linear inequalities, convex quadratic inequalities, matrix norm inequalities, and various constraints from automatic control theory such as Lypaunov inequality as shown above [Boyd *et al.* 1994].

In some cases, the matrix inequality can be a Bilinear Matrix Inequality (BMI), that is a more general case than an LMI.

Definition 10 (**Bilinear Matrix Inequality**, [VanAntwerp & Braatz 2000])**.** *A Bilinear Matrix Inequality (BMI) is of the form:*

$$F(x,y) = F_0 + \sum_{i=1}^{m} x_i F_i + \sum_{j=1}^{l} y_j G_j + \sum_{i=1}^{m}\sum_{j=1}^{l} x_i y_j H_{ij} > 0, \tag{4.3}$$

where $x = [x_1, x_2, \cdots, x_m]^\top \in \mathbb{R}^m$ and $y = [y_1, y_2, \cdots, y_l]^\top \in \mathbb{R}^l$ are the vectors of decision variables; and $F_i \in \mathbb{R}^{n \times n}, \quad \forall i = 0, \cdots, m$, $G_j \in \mathbb{R}^{n \times n}, \quad \forall j = 1, \cdots, l$, $H_{ij} \in \mathbb{R}^{n \times n}, \quad \forall i = 1, \cdots, m$ and $\forall j = 1, \cdots, l$ are given symmetric matrices.

A BMI is an LMI in x for fixed y, and an LMI in y for fixed x; and so is convex in x and convex in y. But, the bilinear terms make the set not jointly convex in x and y [VanAntwerp & Braatz 2000].

In many automatic control problem, it is common to encounter the constraint that some functions be negative whenever some other functions are all negative. In some cases, this constraint can be expressed as an LMI, but in some others we can only approximate this constraint by an LMI, which is more conservative than the initial constraint but is often a useful approximation [VanAntwerp & Braatz 2000].

4.4 Sum Of Squares (SOS)

In this thesis, we are going to rely on a technique called Sum Of Square (SOS), which allows casting several classes of polynomial problems into convex optimization problems. We recall here briefly the basic notions that are necessary for understanding this thesis, the interested reader can then consult the references that are provided.

Definition 11 (SOS problems, [Parrilo 2003]). *Let $p(t) \in \mathbb{R}_{2d}[t]$; we call Sum Of Squares problem (SOS) the problem of finding whether there exists a finite number l of polynomials $\pi_i(t) \in \mathbb{R}_d[t]$ such that*

$$p(t) = \sum_{i=1}^{l} \pi_i(t)^2. \tag{4.4}$$

If such an expression exists, then $p(t)$ is a sum of squares (SOS), which implies that $p(t) \geqslant 0$ for all t.

The class above can be extended to the class of Matrix Sum Of Square (MSOS).

Definition 12 (MSOS problems, [Chesi 2010]). *Let $\mathcal{P}(t) \in \mathbb{R}_{2d}^{n\times n}[t]$; we call a Matrix Sum Of Squares problem (MSOS) the problem of finding whether there exists a finite number l of matrices of polynomials $\Pi_i(t)$ such that*

$$\mathcal{P}(t) = \sum_{i=1}^{l} \Pi_i(t)^\top \Pi_i(t). \tag{4.5}$$

If such a decomposition exists, then $P(t)$ is a matrix sum of squares (MSOS), which implies that $P(t) \succeq 0$ for all t.

The definitions above are restricted to univariate polynomials, as this is the case we consider in this thesis. SOS and MSOS problems are convex problems. A derived class of problems is that of feasibility problems under SOS or MSOS constraints.

Definition 13 (SOS constraints feasibility, [Parrilo 2003]). *Let $p_i(t,\theta) \in \mathbb{R}_{2d}[t]$, $\forall i \in \{1,2,...,q\}$, where $\theta \in \mathbb{R}^\rho$ is a vector of parameters or unknowns, with $p_i(t,\theta)$ affine with respect to the entries of θ. A feasibility problem under SOS constraints consists in finding, if it exists, a value of $\theta = \theta^*$ for which*

$$p_i(t,\theta^*) \text{ is SOS, for } i = 1,...,q. \tag{4.6}$$

If such a θ^* exists, then the problem is feasible; otherwise it is unfeasible.

Definition 14 (MSOS constraints feasibility, [Chesi 2010]). *Let $\mathcal{P}_i(t,\theta) \in \mathbb{R}_{2d}^{n\times n}[t]$, $\forall i \in \{1,2,...,q\}$, where $\theta \in \mathbb{R}^\rho$ is a vector of parameters or unknowns, and the matrices $P_i(t,\theta)$ are affine with respect to the entries of θ. A feasibility problem under SOS constraints consists in finding, if it exists, a value of $\theta = \theta^*$ for which*

$$P_i(t,\theta^*) \text{ is MSOS, for } i = 1,...,q. \tag{4.7}$$

If such a θ^* exists, then the problem is feasible; otherwise it is unfeasible.

Feasibility problems under SOS and MSOS problems are convex optimization problems, and they can be reformulated as linear matrix inequality (LMI) feasibility problems; this can either be done explicitly, or relying on automated procedures, like the one available in the Yalmip toolbox [Löfberg 2009] under Matlab. Minimizing any single affine function of the decision variables under SOS and/or MSOS constraints is also a convex optimization problem.

4.5 The generalized S-procedure

The S-procedure allows restricting some classes of inequalities to a certain given subset of the variables that are concerned [Boyd *et al.* 1994]. In this thesis, we rely on a general expression which can be specialized according to the cases.

Theorem 1. ***S-procedure***
Let $F(x)$, $G(x)$ be (symmetric matrix) functions of the (vector) variable x; and let $g(x)$ be a scalar function of x.
If there exists $\lambda \geq 0$ such that:

$$F(x) - \lambda G(x) \geq 0, \quad \forall x \tag{4.8}$$

Then, the following holds:

$$F(x) \geq 0 \text{ for } G(x) \geq 0, \quad \forall x \tag{4.9}$$

Similarly, if there exists $\Lambda \geq 0$ such that:

$$F(x) - \Lambda g(x) \geq 0, \quad \forall x \tag{4.10}$$

Then, the following holds:

$$F(x) \geq 0 \text{ for } g(x) \geq 0, \quad \forall x \tag{4.11}$$

The terms Λ and λ are called multipliers, which can be chosen at one's convenience, i.e. they are decision variables subject to the positivity constraints above. When used in the context of polynomial problems as in this case, the lemma above is a direct consequence of a lemma known as Positivstellensatz [Chesi 2010] or p-satz, of which several versions exist in the literature. The multipliers can in this case have a polynomial dependence on x, which allows satisfying the positivity constraints by means of either SOS or MSOS constraints.

4.6 Set-theory preliminaries

Set theoretic-based methods are methods utilizing certain type of set to model a control problem (e.g. uncertainty, target objective, obstacle) [Blanchini & Miani 2015]. In particular, they exploit the structural property of convexity of functions and sets to formulate their problem as convex optimization problem. Main convex sets encountered are: interval set, polyhedral set, zonotopic set, and ellipsoidal set [Blanchini 1999]. The set theoretic-based methods developed in this thesis are based on ellipsoidal sets and halfspaces, so their mathematical form will be detailed. First and foremost, some basic set definitions and operations are given.

4.6.1 Convex functions and sets

Definition 15 (**Convex set**, [Lucchetti 2006]). *A set $C \subseteq \mathbb{R}^n$ is convex if, for all x, $y \in C$ and for all $\lambda \in [0, 1]$,*

$$\lambda x + (1 - \lambda) y \in C \tag{4.12}$$

This means that if any two elements in C are taken, and a line segment is drawn between these two elements, then every points on that line segment also belongs to C. Note that the point $\lambda x + (1-\lambda)y$ is called a convex combination of the points x and y.

The definition of convexity for functions is given below.

Definition 16 (**Convex function**, [Lucchetti 2006]). *A function $f : \mathcal{D} \to \mathbb{R}$ is convex if its domain $\mathcal{D} \subseteq \mathbb{R}^n$ is a convex set and if for all x, $y \in$ dom(f) and for all $\lambda \in [0,1]$ the following holds:*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \tag{4.13}$$

This definition can be interpreted geometrically as follows: the line segment connecting $(x, f(x))$ to $(y, f(y))$ should sit above the graph of the function.

There are many attractive properties associated with convexity [Boyd & Vandenberghe 2004]. Among others, a fundamental property given as follows leads many problems to be solved with convex optimization.

Theorem 2. ***(Global minimum).*** *Consider an optimization problem of the form:*

$$\underset{x \in S}{\text{minimize}} \quad f(x) \tag{4.14}$$

where f is a convex function and S is a convex set. Then, any local minimum is also a global minimum.

4.6.2 κ-sublevel sets

[**κ-sublevel set**, [Boyd & Vandenberghe 2004]] Convex functions give rise to an interesting type of convex set called a κ-sublevel set.

Definition 17. ***(κ-sublevel sets).*** *Given a convex function $f : \mathbb{R}^n \to \mathbb{R}$ and a real number $\kappa \in \mathbb{R}$, the κ-sublevel set is defined as*

$$\{x \in \mathcal{D}(f) | f(x) \leq \kappa\} \tag{4.15}$$

In other words, the κ-sublevel set is the set of all points x such that $f(x) \leq \kappa$.

Theorem 3. *If a function $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then all its sublevel sets are convex sets.*

Definition 18 (**κ-level set**, [Boyd & Vandenberghe 2004]). *Given a convex function $f : \mathbb{R}^n \to \mathbb{R}$ and a real number $\kappa \in \mathbb{R}$, the κ-level set is defined as*

$$\{x \in \mathcal{D}(f) | f(x) = \kappa\} \tag{4.16}$$

4.6.3 Operations between sets

Some basic operations between sets are necessary to understand the methods developed in this thesis. They are presented in the following.

Definition 19. ***(Inclusion).*** *A set X is included in a set Y, i.e. $X \subseteq Y$, if and only if $x \in Y$, $\forall x \in X$. This means that X is a subset of Y.*

Definition 20. ***(Intersection).*** *The intersection of two sets X and Y is defined as follows:*

$$X \cap Y = \{z| \quad z \in X, \quad z \in Y\} \tag{4.17}$$

Definition 21. ***(Scaling).*** *The scaling of a set X by a real $\lambda \in R_+$ denoted λX is defined as follows:*

$$\lambda X = \{z = \lambda x| \quad x \in X\}, \quad \lambda \geq 0 \tag{4.18}$$

Some set operations preserve the convexity:

Theorem 4. *Consider the convex set S_1 scaled by a real λ, the scaled set is also convex, i.e. λS_1 is convex.*

4.6.4 Ellipsoidal sets

Ellipsoidal sets also called ellipsoids is a class of convex sets widely used in the automatic control of dynamical systems.

Definition 22. ***[Ellipsoid**, [Blanchini & Miani 2015]] Given a strictly positive definite matrix $Q \in \mathbb{R}^{n \times n}$, i.e. $Q > 0$ and $Q = Q^\top$, a real vector $\bar{x} \in \mathbb{R}^n$, and a strictly positive real scalar ρ, the ellipsoid $\mathcal{E}(Q, \bar{x}, \rho)$ is defined by the set:*

$$\mathcal{E}(Q, \bar{x}, \rho) = \{x \in \mathbb{R}^n| \quad (x - \bar{x})^\top Q(x - \bar{x}) \leq \rho\} \tag{4.19}$$

where Q is the shape matrix of the ellipsoid, $\bar{x}$ its center, and ρ its radius.

A normalized ellipsoid defined in (4.20) is often utilized instead of the general form in 4.19.

$$\mathcal{E}(Q, \bar{x}, 1) = \mathcal{E}(Q, \bar{x}) = \{x \in \mathbb{R}^n| \quad (x - \bar{x})^\top Q(x - \bar{x}) \leq 1\} \tag{4.20}$$

where Q is the shape matrix of the ellipsoid, and $\bar{x}$ its center.

From now on, the term ρ is dropped out in the notation when we refer to a normalized ellipsoid, as shown in (4.20).

Normalized ellipsoids can also be rewritten with $\tilde{x} = [x^\top, 1]^\top$ in the following quadratic form:

$$\mathcal{E}(Q, \bar{x}) = \{\tilde{x} \in \mathbb{R}^{n+1} \,|\, \tilde{x}^\top \tilde{Q} \tilde{x} \leq 1\} \tag{4.21}$$

where

$$\tilde{Q} = \begin{bmatrix} Q & -Q\bar{x} \\ -\bar{x}^\top Q & \bar{x}^\top Q \bar{x} \end{bmatrix} \in \mathrm{R}^{(n+1)\times(n+1)}. \tag{4.22}$$

Note that if the center is set at zero, i.e. $\bar{x} = 0$, then from the definition in 22 the quadratic form becomes:

$$\mathcal{E}(Q, 0) = \mathcal{E}(Q) = \{x \in \mathbb{R}^n| \quad x^\top Q x \leq 1\} \tag{4.23}$$

From now on, the term $\bar{x}$ is dropped out in the notation when the center is null, as shown in (4.23).

4.6.5 Hyperplanes and halfspaces

Definition 23. *[**Hyperplane**, [Kurzhanskiy & Varaiya 2006]] A hyperplane denoted $\mathcal{H}(c,\gamma) \in \mathbb{R}^n$ is the set defined as follows:*

$$\mathcal{H}(c,\gamma) = \{x \in \mathbb{R}^n | \quad c^\top x = \gamma\} \tag{4.24}$$

where $c \neq \mathbb{O} \in \mathbb{R}^n$, and $\gamma \in \mathbb{R}$ are given.

A hyperplane defines two (closed) halfspaces:

$$S_1 = \{x \in \mathbb{R}^n | \quad c^\top x \leq \gamma\} \tag{4.25}$$

and

$$S_2 = \{x \in \mathbb{R}^n | \quad c^\top x \geq \gamma\} \tag{4.26}$$

To avoid confusion, we assume that a hyperplane $\mathcal{H}(c,\gamma)$ specifies the halfspace in (4.26). The other halfspace defined in (4.25) is never used in the remaining of this manuscript.

4.6.6 Distance between an ellipsoid and an hyperplane

Definition 24. *[**Distance function**, [Kurzhanskiy & Varaiya 2006]] Consider a (normalized) ellipsoid $\mathcal{E}(Q,\bar{q})$ and a hyperplane $\mathcal{H}(c,\gamma)$. The distance from $\mathcal{E}(Q,\bar{q})$ to $\mathcal{H}(c,\gamma)$ is given as follows:*

$$\text{dist}(\mathcal{E}(Q,\bar{q}), \mathcal{H}(c,\gamma)) = \frac{(c^\top Q^{-1} c)^{1/2} - |\gamma - c^\top \bar{q}|}{(c^\top c)^{1/2}} \tag{4.27}$$

where $\text{dist}(\mathcal{E}(Q,\bar{q}), \mathcal{H}(c,\gamma)) \geqslant 0$ if the ellipsoid intersects the hyperplane; otherwise it is negative.

Note that in the case when the distance function is equals to zero, this means that the hyperplane is a supporting hyperplane for the ellipsoid.

4.7 Convex optimization problems

In this thesis, the proposed methods are based on convex optimization problems. From the definitions of convex functions and convex set, we present the convex optimization problems in this Section [Boyd & Vandenberghe 2004].

Definition 25. *[**Convex optimization problem**] Formally, a convex optimization problem is an optimization problem of the form*

$$\begin{aligned} &\underset{x}{\text{minimize}} && f(x) \\ &\text{subject to} && x \in C \end{aligned}$$

where f is a convex function defining the cost function, , $x \in \mathbb{R}^m$ is the vector of decision variables, and C is a convex set. From this formal definition of convex optimization problem, we often rewrite is as

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \quad i = 1, ..., m \\ & h_i(x) = 0, \quad i = 1, ..., p \end{aligned} \tag{4.28}$$

where f is a convex function defining the cost function, , g_i are convex functions, and h_i are affine functions, and $x \in \mathbb{R}^m$ is the vector of decision variables.

Remark 2. *The direction of the inequalities in convex optimization problem of the form (4.28) matters: a convex function g_i must be less than zero. Indeed the 0-sublevel set of g_i is a convex set, thus the feasible region, i.e. the intersection of many convex sets defined by g_i, is also convex. If the constraint becomes $g_i \geq 0$ for some convex function g_i, then we are interested in 0-superlevel set, not 0-sublevel set. In this case, the feasible region is no longer a convex set. Note that superlevel sets are convex for concave functions, i.e. if f is a concave function and $g_i \geq 0$ then the feasible region is convex.*

Remark 3. *Only affine functions are allowed to be equality constraints. Indeed, an equality constraint $h_i = 0$ is equivalent to two inequality constraints $h_i \geq 0$ and $h_i \leq 0$. However, this is only valid constraints if and only if h_i is both convex and concave, i.e. h_i must be affine.*

Before stating the main property of global optimality in convex optimization problems, the concepts of local optima and global optima are defined.

Definition 26. ***(Locally optimal point).*** *A point x is locally optimal if it is feasible, i.e. it satisfies the constraints of the optimization problem, and if there exists some $R > 0$ such that all feasible points z with $||x - z||_2 \leq R$ satisfy $f(x) \leq f(z)$*

where $||.||_2$ is the Euclidean norm.

Definition 27. ***(Globally optimal point).*** *A point x is globally optimal if it is feasible, i.e. it satisfies the constraints of the optimization problem, and for all feasible points z, $f(x) \leq f(z)$*

One of the main property in convex optimization problems is stated as follows.

Property 1. ***(Global optimality).*** *For a convex optimization problem, all locally optimal points are globally optimal.*

As a final note, convex optimization problems are often considered as the largest class of optimization problems for solving engineering problems [Boyd & Vandenberghe 2004]. A short list of the main reasons is given below.

1. Broadest class of optimization problems one can solve efficiently [Boyd & Vandenberghe 2004];

2. Global optimality is an appealing property [Boyd & Vandenberghe 2004];

3. Automated procedures to solve them are available (e.g. Yalmip toolbox under Matlab) [Lofberg 2004];

4. Many solvers exist (e.g. Mosek, cvx) [Aps 2020, Boyd & Grant 2020];

5. Many engineering problems can be cast into convex optimization problems.

Note that solving convex optimization problem is often referred as convex programming.

Remark 4. *All the convex optimization problems appearing in the methods presented later in this thesis have been solved with Mosek solver with Yalmip toolbox on Matlab R2019a.*

4.7.1 Special cases of convex optimization problems

After having presented the general convex optimization problems, some special cases used in this thesis are presented in the following.

1. **Linear Programming (LP)** is a convex optimization problem in which both the cost function f and inequality constraints g_i are affine functions. Hence, these problems have the form

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^\top x + d \\ \text{subject to} & Gx \leq h, \qquad Ax = b \end{array} \tag{4.29}$$

   where $x \in \mathbb{R}^n$ is the vector of decision variables, and $c \in \mathbb{R}^n$, $d \in \mathbb{R}$, $G \in \mathbb{R}^{m \times n}$, $h \in \mathbb{R}^m$, $A \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^p$ are defined by the problem. Here, the operator '$\geq$' denotes elementwise inequality, not semidefiniteness.

2. **Semidefinite Programming (SDP)** is a convex optimization problem if it is of the form

$$\begin{array}{ll} \underset{X}{\text{minimize}} & \text{Trace}(CX) \\ \text{subject to} & \text{Trace}(A_i X) = b_i, \quad i = 1, ..., p \\ & X \geq 0 \end{array} \tag{4.30}$$

   where the symmetric matrix $X \in \mathbb{R}^{n \times n}$ is the decision variable, the symmetric matrices $C, A_1, \ddots, A_p \in \mathbb{R}^{n \times n}$ are defined by the problem, and the constraint $X \geq 0$ means that we are constraining X to be positive semidefinite, i.e. it is an LMI.

To conclude, all these sections have presented the sufficient background for understanding the rest of this manuscript. In addition to this, Lyapunov functions are explained and detailed in the next section. This is a fundamental background for our approach.

4.8 Lyapunov function

Lyapunov functions is a fundamental concept in the theory of stability of dynamical systems. It is utilized in the famous second method of Lyapunov [Kalman & Bertram 1960, Lyashevskiy & Meyer 1995]. The principal idea of the second method is contained in the following physical reasoning: "if the rate of change $\frac{dE(x)}{dt}$ of the energy $E(x)$ of an isolated physical system is negative for every possible state x, except for a single equilibrium state x_e, then the energy will continually decrease until it finally assumes its minimum value $E(x_e)$" [Kalman & Bertram 1960]. This can be interpreted as follows: a system with this property perturbed from its equilibrium state will always return to it. This is the so called concept of Lyapunov stability. Unfortunately, there is no straightforward method to define the energy $E(x)$ for a dynamical system, but the mathematical counterpart of the foregoing reasoning is the following theorem:

Definition 28 (**Equilibrium state**, [Kalman & Bertram 1960])**.** *The state x_e of a system is the equilibrium state of a dynamical system as in (3.3) if and only if the time derivative of x is zero, i.e. $\dot{x}(t) = 0$.*

Remark 5. *For the autonomous system, the equilibrium state x_e is expressed as*

$$0 = Ax_e + a \tag{4.31}$$

For the case of a system with a null affine term, it is trivial to show that $x_e = 0$; and $x_e \neq 0$ for $a \neq 0$.

Definition 29 (**Laypunov function**, [Kalman & Bertram 1960])**.** *A Lyapunov function is any scalar function $V(\tilde{x})$ of the state with the properties:*

1. *$V(\tilde{x}) \geq 0$, $\dot{V}(\tilde{x}) \leq 0$ when $x(t) \neq x_e$*
2. *$V(\tilde{x}) = \dot{V}(\tilde{x}) = 0$ when $x(t) = x_e$*

Theorem 5 (**Lyapunov stability**, [Kalman & Bertram 1960])**.** *A dynamical system as in (3.3) is Lyapunov stable if and only if there exists a Lyapunov function.*

In this case we say that the system is Lyapunov stable, and its state returns to the equilibrium after a perturbation.

Remark 6. *Unlike the energy $E(x)$ of a system, the Lyapunov function $V(x)$ is not unique. The energy $E(x)$ is not necessary a Lyapunov function: a system whose energy $E(x)$ decreases on the average, but not at each instant, is stable but $E(x)$ is not a Lyapunov function as the requirement $\dot{V}(x) < 0$, $\forall x \neq x_e$ is not satisfied at each instant.*

The definition of a Lyapunov function and the Theorem above are usually stated for the special case where the affine term is null with $V(x) = x^\top Q x > 0$ when $x \neq 0$. This can be shown as follows: consider the system in (3.3) and $V(\tilde{x}) = \tilde{x}(t)^\top \tilde{Q} \tilde{x}(t)$ with $\tilde{Q} \geq 0$ a Lyapunov function. From , recall that

$$V(\tilde{x}) = \tilde{x}(t)^\top \tilde{Q} \tilde{x}(t) \leftrightarrow V(x) = (x - \bar{x})^\top Q (x - \bar{x}) \tag{4.32}$$

where $Q > 0$, which can be rewritten $(x - \bar{x})^\top Q(x - \bar{x}) > 0$ when $x \neq x_e$. As $V(x)$ is a Lyapunov function, then $V(x) = 0$ when $x = x_e$. Hence it is the case only if $\bar{x} = x_e$. Recalling that the equilibrium state $x_e = 0$ for system as in (3.3) with null affine term. Thus, by setting $\bar{x} = x_e = 0$ in (4.33) it yields

$$V(x) = x^\top Q x \tag{4.33}$$

where $Q > 0$, which can be rewritten $x^\top Q x > 0$ when $x \neq 0$. This implies that $V(x) = x^\top Q x$ is a Lyapunov function for the system in (3.3) with null affine term.

The concepts presented in this section are illustrated by the following two examples.

Example 1. *Consider an harmonic oscillator described by the linear differential equation*

$$\ddot{x_1} + x_1 = 0 \tag{4.34}$$

or equivalently in the state-space representation as in (3.3) with $x = [x_1, x_2]^\top$ where $\dot{x_1} = x_2$, and $a = 0$

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{4.35}$$

The trajectories of (4.34) or equivalently of (4.35) in the state-plane are on the boundaries of ellipsoids about the origin, i.e. $x_e = 0$. The energy of the system is given by

$$E(x) = x_1^2 + x_2^2 = V(x) \tag{4.36}$$

The time derivative $\dot{V}(x)$ of $V(x)$ along any trajectory of (4.34) or equivalenty (4.35) is given by

$$\begin{aligned} \dot{V}(x) &= \dot{x}'[\nabla V(x)] \\ &= \dot{x}_1 \frac{\partial V}{\partial x_1} + \dot{x}_2 \frac{\partial V}{\partial x_2} \\ &= 2\dot{x}_1 x_1 + 2\dot{x}_2 x_2 \end{aligned} \tag{4.37}$$

By replacing (4.34) in (4.37), it yields $\dot{V}(x) = 0$, so that $V(x) > 0$ and $\dot{V}(x) = 0$. It can be interpreted physically as follows: the energy remains constant, i.e. or is a conservative system. Thus, the state trajectories never return to the equilibrium state. Hence, the dynamical system is not Lyapunov stable.

Example 2. *Consider now a slight modification of (4.34) in the state-space representation:*

$$A = \begin{bmatrix} -c_1 & 1 \\ -1 & -c_2 \end{bmatrix} \tag{4.38}$$

with c_1, c_2 two fixed positive constants. Let define $V(x)$ as in (4.36), but now this quantity is not necessarily the energy of system in (4.38). Calculating as before, the derivative $\dot{V}(x)$ of $V(x)$ along the trajectory of (4.38) is:

$$\dot{V}(x) = -2c_1 x_1^2 - 2c_2 x_2^2 < 0, \text{ when } x \neq 0 \tag{4.39}$$

Hence, $V(x) > 0$ and $\dot{V}(x) < 0$ when $x \neq 0$. This implies that $V(x)$ is a Lyapunov function for system (4.38); so the system is Lyapunov stable as a Lyapunov function exists.

Notice that the state trajectory crosses the boundary ellipsoid defined by the κ-sublevel set from the outside toward the inside until reaching the origin, i.e. the origin is asymptotically stable.

As a final remark, the problem of existence of a quadratic Lyapunov function for a dynamical system as in (3.3) can be cast into a SDP feasability problem with LMI constraints.

$$\begin{cases} \text{Find } \tilde{Q} \text{ such that} \\ V(\tilde{x}) \geq 0 \\ \dot{V}(\tilde{x}) \leq 0 \end{cases} \Leftrightarrow \begin{cases} \text{Find } \tilde{Q} \text{ such that} \\ \tilde{Q} \geq 0 \\ -\tilde{A}^\top \tilde{Q} - \tilde{Q}\tilde{A} \geq 0 \end{cases} \tag{4.40}$$

with

$$\dot{V}(\tilde{x}) = \dot{\tilde{x}}^\top \tilde{Q}\tilde{x} + \tilde{x}^\top \tilde{Q}\dot{\tilde{x}} \tag{4.41}$$

$$= \tilde{x}^\top (\tilde{A}^\top \tilde{Q} + \tilde{Q}\tilde{A})\tilde{x} \tag{4.42}$$

4.9 Conclusion

This Chapter has presented the sufficient background for understanding the methods developed in this thesis. In the following Chapters we will detail the methods of this thesis for analyzing and preventing stealthy aging attacks.

Chapter 5

Analysis and prevention: a subcase of stealthy aging attacks

Contents

5.1 Introduction

As introduced in Chapter 3, our methods contribute in the analysis and prevention of stealthy aging attacks. Among these attacks, two of them manipulating the control signal are considered in this thesis: (AT1) where the objective is to let the dynamical system reach a subset of the state space featuring an abnormal aging; and (AT2) where the objective is to maximize the time integral of a degradation function modeling the stresses acting on the process. As we will see in the following Chapters, AT1 is a subcase of AT2. This Chapter is devoted to the subcase of stealthy aging attacks (AT1).

Section 5.2 formally defines the subcase of stealthy aging attacks (AT1) to provide the framework of the problem the methods will solve. Thereafter, the methods for analyzing and preventing stealthy aging attacks of AT1 are detailed in Section 5.3 and Section 5.4.

5.2 Research question

From the general definition of stealthy aging attacks in Definition 5, we restrain our investigation to the following Attack Type (AT).

Definition 30. ***(Attack Type 1 (AT1)).*** *The input $u(t)$ is manipulated by an attacker to let the dynamical system reach a subset D_x of the state space that we call abnormal aging set, because it features an aging faster than normal for the dynamical system, leading to lifetime reduction. In the meantime, the state trajectory is constrained to remain in a set $\mathcal{X}_s$ we call service set, because this is the set the state trajectory visits during the normal operation of the control system, i.e. the attack-free case.*

As an interpretation of the Definition 30, if the state trajectory $x(t)$ of the dynamical system reaches the abnormal aging set D_x then the dynamical system operates outside of the normal aging conditions. In the meantime, the state trajectory remains in the service set X_s, where the delivered service is normal. If the state trajectory remains inside the service set, we can consider that the service can be provided normally, such that no suspicion of attack can arise. Clearly, AT1 is possible only if $X_s \cap D_x$ is non-empty, otherwise stealthy aging attacks of AT1 are not feasible.

The general idea to analyze and prevent such kind of attacks is summarized as follows. Consider a dynamical system as in (3.3) where $u(t)$ is bounded at every time instant due to physical input limitation (e.g. voltage, current, torque):

$$u(t) \in \mathcal{U}, \quad \forall t \geq 0 \tag{5.1}$$

where $\mathcal{U}$ is called the control input set, and it is a closed set.

Note that we insist in the physical nature of the input limitation as it constrains the attack capabilities. In fact, virtual limitation imposed by the controller could be bypassed by an attacker compromising the controller. It is important to note that no other conditions are imposed on the input $u(t)$ (e.g. band-limited signal).

Consider $\mathcal{X}$ the state set of the dynamical system such that $x(t) \in \mathcal{X}, \quad \forall t \geq 0$. As the attacker wants to remain stealthy regarding the delivered service, we may assume that the control signal $u(t)$ will be chosen carefully in order to not degrade the service delivered by the dynamical system. For this reason, we can assume that a part of the state can be constrained to belong to $\mathcal{X}_s$, the set that will not cause any suspicions of attack to arise:

$$x(t) \in \mathcal{X} \subseteq \mathcal{X}_s, \quad \forall t \geq 0 \tag{5.2}$$

As a consequence of this, the following assumption is made:

Assumption 1. *There exists a control signal $u(t) \in \mathcal{U}$ such that $x(t) \in \mathcal{X}_s$ for all $t \geq 0$ and for any initial conditions $x(t_0) \in \mathcal{X}$.*

If the assumption above is not satisfied, it means that no stealthy aging attack can be performed without being detected from the sensors measuring the delivered service, which put us on the safe side. This assumption will be refined later to be more specific for each method, but the essence of the assumption remains.

From this problem, the following two Research Question (RQ) are addressed in this Chapter:

Definition 31. ***(RQ1).*** *Are stealthy aging attacks of AT1 feasible for the given dynamical system in (3.3), the control input set $\mathcal{U}$, the service set $\mathcal{X}_s$, and the abnormal aging set D_x?*

Definition 32. ***(RQ2).*** *How the control input set $\mathcal{U}$ can be restrained to $\hat{\mathcal{U}}$ such that stealthy aging attacks of AT1 are not feasible, i.e. the answer to RQ1 is negative?*

After having defined the Attack Type of stealthy aging attacks we consider, stated the problem, and defined the research questions we answer in this Chapter, the framework of the developed methods for analyzing and preventing AT1 is presented.

5.2.1 Framework of the methods

From the physical input limitation $\mathcal{U}$ in (5.1), the proposed methods perform the following steps:

1. Computation of the reachable set under the given constraints, i.e. $\mathcal{X}$. This analysis step consists in quantifying the effect induced by AT1 on the dynamical system.

2. Reachability analysis of the abnormal aging set D_x. This consists in assessing if the state trajectory can reach the abnormal aging set D_x, i.e. $\mathcal{X} \cap D_x \neq \emptyset$. In the case of $\mathcal{X} \cap D_x = \emptyset$, it means that the state trajectory cannot reach the abnormal aging set, so stealthy aging attacks of AT1 are not feasible (RQ1); hence (RQ2) is already answered: the current physical input limitation $\mathcal{U}$ is sufficient to prevent stealthy aging attacks. Otherwise, stealthy aging attacks of AT1 are feasible (RQ1), so (RQ2) is addressed with step 3.

3. Restraining the control input set $\mathcal{U}$ to $\hat{\mathcal{U}}$ to make (RQ1) not feasible, with $\hat{\mathcal{U}} \subseteq \mathcal{U}$. It consists in finding the largest control input set $\hat{\mathcal{U}}$ such that the constrained state set $\mathcal{X}$ does not intersect with the abnormal aging set D_x, i.e. $\mathcal{X} \cap D_x = \emptyset$ if $u(t) \in \hat{\mathcal{U}}, \quad \forall t \geq 0$. This answers RQ2.

5.2.2 Modeling considerations

In order to make the problem solvable with convex optimization problems, the following considerations are made:

- The control input set $\mathcal{U}$ is modeled as an ellipsoid $\mathcal{E}_u(R, \bar{u})$, i.e. $u(t) \in \mathcal{U} \subseteq \mathcal{E}_u$, $\forall t \geq 0$.
- The constrained control input set $\hat{\mathcal{U}}$ is modeled as an ellipsoid $\hat{\mathcal{E}}_u$, i.e. $u(t) \in \hat{\mathcal{U}} \subseteq \hat{\mathcal{E}}_u$, $\forall t \geq 0$.
- The state set $\mathcal{X}$ is modeled as an ellipsoid $\mathcal{E}_x(Q, \bar{x})$, i.e. $x(t) \in \mathcal{X} \subseteq \mathcal{E}_x$, $\forall t \geq 0$.
- The service set $\mathcal{X}_s$ is modeled as an ellipsoid $\mathcal{E}_s(\Xi, \bar{\xi})$, i.e. $x(t) \in \mathcal{X} \subseteq \mathcal{E}_x$, $\forall t \geq 0$.

- The abnormal aging set $\mathcal{D}_x$ is modeled as the union of i number of halfspaces $\mathcal{H}_i(c_i, \mu_i)$ defined by their boundary hyperplane:

$$\mathcal{D}_x = \{x \in \mathbb{R}^n : \bigcup_{i=1}^{p} c_i^\top x(t) \geq \mu_i\} \tag{5.3}$$

From the modeling considerations and before detailing the methods proposed in this thesis, we express the constraints (5.1) and (5.2):

$$u(t) \in \mathcal{E}_u(R, \bar{u}) \Leftrightarrow \quad (u - \bar{u})^\top R(u - \bar{u}) \leq 1 \tag{5.4}$$

$$x(t) \in \mathcal{E}_s(\Xi, \bar{\xi}) \Leftrightarrow \quad (x - \bar{\xi})^\top \Xi(x - \bar{\xi}) \leq 1 \tag{5.5}$$

Before reformulating the constraints, let

$$\Gamma^\top = \begin{bmatrix} 0_{n,1} \\ 1 \end{bmatrix} \tilde{\Xi} = \begin{bmatrix} \Xi & 0_{n,1} \\ 0_{1,n} & 0, \end{bmatrix}, \tilde{\xi} = \begin{bmatrix} \bar{\xi} \\ 0 \end{bmatrix}. \tag{5.6}$$

By introducing the vector $\begin{bmatrix} \tilde{x} \\ u \end{bmatrix}$, it can be restated as follows.

$$(u - \bar{u})^\top R(u - \bar{u}) \leq 1 \Leftrightarrow \begin{bmatrix} \tilde{x} \\ u \end{bmatrix}^\top \begin{bmatrix} \Gamma^\top \bar{u}^\top R \bar{u} \Gamma & -\Gamma^\top \bar{u}^\top R \\ -R\bar{u}\Gamma & R \end{bmatrix} \begin{bmatrix} \tilde{x} \\ u \end{bmatrix} \leq 1 \tag{5.7}$$

$$(x - \bar{\xi})^\top \Xi(u - \bar{\xi}) \leq 1 \Leftrightarrow \begin{bmatrix} \tilde{x} \\ u \end{bmatrix}^\top \begin{bmatrix} \tilde{\Xi} - \tilde{\Xi}\tilde{\xi}\Gamma - \Gamma^\top \tilde{\xi}^\top \tilde{\Xi} + \Gamma^\top \tilde{\xi}^\top \tilde{\Xi} \tilde{\xi} \Gamma & 0_{n+1,m} \\ 0_{m,n+1} & 0_{m,m} \end{bmatrix} \begin{bmatrix} \tilde{x} \\ u \end{bmatrix} \leq 1 \tag{5.8}$$

Rewriting the matrix inequalities in (5.7) in the positive semidefinite form (≥ 0), it yields:

$$\begin{bmatrix} \tilde{x} \\ u \end{bmatrix}^\top \begin{bmatrix} -\Gamma^\top \bar{u}^\top R \bar{u} \Gamma & -\Gamma^\top \bar{u}^\top R \\ -R\bar{u}\Gamma & R \end{bmatrix} \begin{bmatrix} \tilde{x} \\ u \end{bmatrix} + 1 \geq 0, \tag{5.9}$$

$$\begin{bmatrix} \tilde{x} \\ u \end{bmatrix}^\top \begin{bmatrix} -\tilde{\Xi} + \tilde{\Xi}\tilde{\xi}\Gamma + \Gamma^\top \tilde{\xi}^\top \tilde{\Xi} - \Gamma^\top \tilde{\xi}^\top \tilde{\Xi} \tilde{\xi} \Gamma & 0_{n+1,m} \\ 0_{m,n+1} & 0_{m,m} \end{bmatrix} \begin{bmatrix} \tilde{x} \\ u \end{bmatrix} + 1 \geq 0 \tag{5.10}$$

Note that $\Gamma^\top \Gamma = \begin{bmatrix} 0_{n,n} & 0_{n,1} \\ 0_{1,n} & 1 \end{bmatrix}$, so that $\tilde{x}^\top \Gamma^\top \Gamma \tilde{x} = \begin{bmatrix} x \\ 1 \end{bmatrix}^\top \begin{bmatrix} 0_{n,n} & 0_{n,1} \\ 0_{1,n} & 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}$ allows the expression of the constant term. Therefore, the constraints can be restated as follows:

$$u(t) \in \mathcal{E}_u(R, \bar{u}) \Leftrightarrow \quad S \geq 0 \tag{5.11}$$

$$x(t) \in \mathcal{E}_s(\Xi, \bar{\xi}) \Leftrightarrow \quad T \geq 0 \tag{5.12}$$

with

$$S = \begin{bmatrix} \Gamma^\top \Gamma - \Gamma^\top \bar{u}^\top R \bar{u} \Gamma & \Gamma^\top \bar{u}^\top R \\ R\bar{u}\Gamma & -R \end{bmatrix}$$

$$T = \begin{bmatrix} \Gamma^\top \Gamma - \tilde{\Xi} + \tilde{\Xi}\tilde{\xi}\Gamma + \Gamma^\top \tilde{\xi}^\top \tilde{\Xi} - \Gamma^\top \tilde{\xi}^\top \tilde{\Xi} \tilde{\xi} \Gamma & 0_{n+1,m} \\ 0_{m,n+1} & 0_{m,m} \end{bmatrix}$$

The service set is defined as $\mathcal{E}_x(\Xi, \bar{\xi})$, where Ξ in general will be rank-deficient, as it will constrain only one part of the state, that is involved in the service to deliver by the physical system; $\mathcal{E}_x(\Xi, \bar{\xi})$ can even coincide with $\mathbb{R}^{n \times n}$ (formally by picking $\Xi = 0$).

Remark 7. *If $\Xi = 0$ then the state trajectory is not constrained to belong to the service set. This is the case of an open-loop system where the delivered service is not measured by sensors neither operators.*

After having expressed the constraints $u(t) \in \mathcal{E}_u$ and $x(t) \in \mathcal{E}_s$, $\forall t \geq 0$, we present in the next sections the two set theoretic methods we propose. The first one called invariant set-based method is based on the invariance concept, while the second one called convex robust simulation-based method is based on the simulation of a dynamical system to forecast its state set. Beforehand, the main steps of the proposed methods are pictured in the following figures.

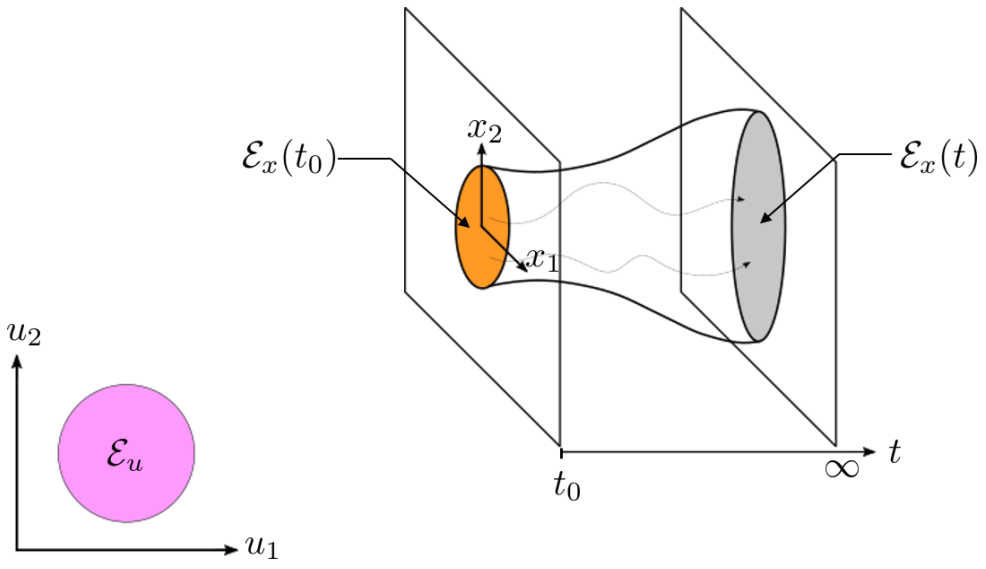


Figure 5.1: Step 1 : Analysis of stealthy aging attacks

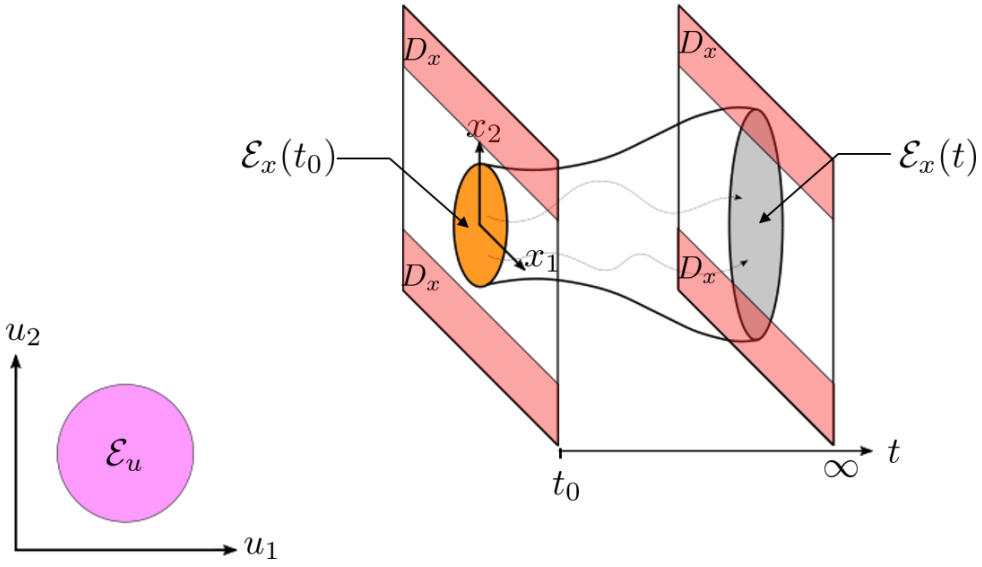


Figure 5.2: Step 2 : Reachability analysis of the abnormal aging set

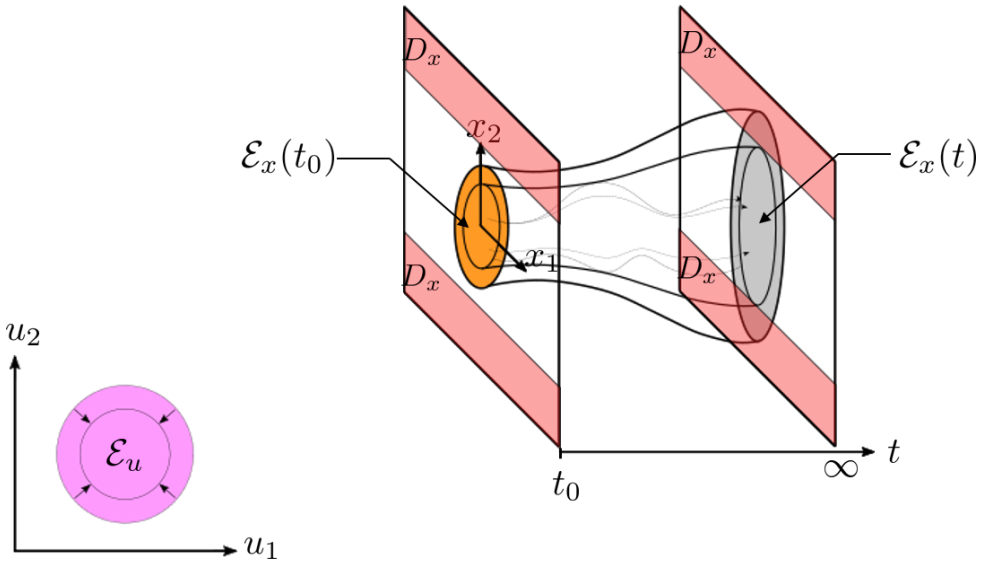


Figure 5.3: Step 3 : Restriction of the input set to prevent such attacks

5.3 Invariant set-based method

5.3.1 Invariance notions

In this method [Escudero & Zamaï 2019], we propose to exploit the property of positive invariance of an ellipsoid, defined as follows.

Definition 33 (**Positive invariance**, [Gayek 1991]). *An ellipsoid $\mathcal{E}_x$ is positive invariant for a dynamical system as in (3.3) if from any initial conditions $x(0) \in \mathcal{E}_x$ and for all $u(t) \in \mathcal{E}_u$, $x(t) \in \mathcal{E}_x$, $\forall t \geq 0$.*

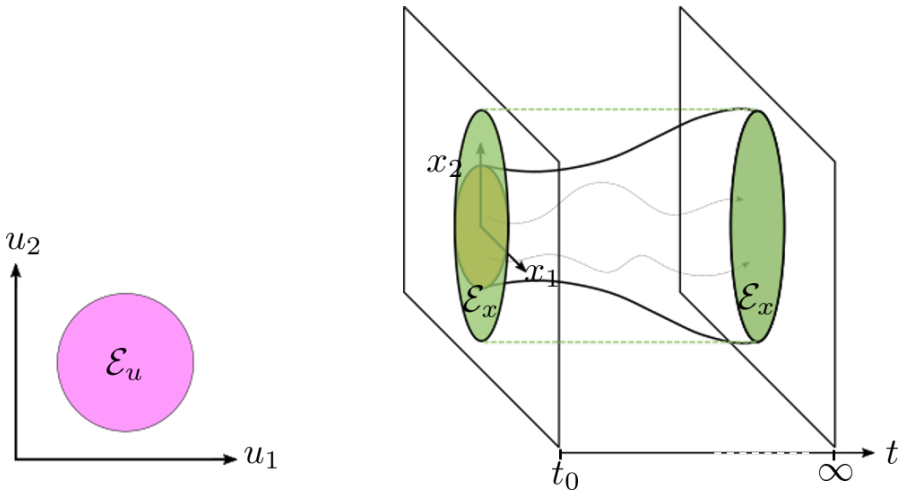


Figure 5.4: Invariant ellipsoid $\mathcal{E}_x$

Remark 8. *The role of the term "positive" is referred to the fact that the invariant property regards the future. This term is commonly dropped and the ellipsoid is said to be an invariant set or an invariant ellipsoid.*

The property of invariance is then attractive: it allows the encapsulation of the state set at the infinite-time horizon of the dynamical system, i.e. $\mathcal{X}(\infty, x(0)) \subseteq \mathcal{E}_x$ where $\mathcal{E}_x$ is invariant. Hence, an invariant ellipsoid is an outer-approximation of $\mathcal{X}(\infty, x(0))$. This set $\mathcal{X}(\infty, x(0))$ often called the infinite time reachable set of the dynamical system states is the set that captures all the state trajectory emanating from initial states $x(0)$. However, its analytic computation is generally not possible. Instead, invariant sets are computed as they benefit from the fact that their family can be chosen at one convenience. The choice of the set family has an effect on the computational complexity, and the error of approximation of the reachable set.

The positive invariance notion has largely been utilized in the analysis and synthesis of controllers for dynamical systems. In particular, two relevant sets in the automatic control can be pinpointed [Blanchini & Miani 2015]: the robustly (positively) invariant set, and the robust controlled (positively) invariant set. The robustly invariant set is the set that encompasses the state trajectory while the dynamical system is subject to disturbance. The robustly controlled invariant set is the set that encompasses the states trajectory while the dynamical system is subject to disturbance and a control input computed from a controller.

5.3.2 Main theorem of the invariant set-based method

The main theorem of this method provides sufficient conditions for computing the invariant state ellipsoid for a dynamical system as in (3.3) with the state trajectory constrained to belong to the service set. This is what we call the constrained invariant state ellipsoid. This theorem will be subsequently employed in a set of algorithms, whose goal is the prediction and prevention of potential stealthy aging attack of AT1.

This method is based on the search of a pseudo-Lyapunov function $V(\tilde{x}) = \tilde{x}^\top \tilde{Q}\tilde{x}$. The idea is to characterize a positive definite function

$$V(\tilde{x}) = \tilde{x}^\top \tilde{Q}\tilde{x} \geq 0 \ \forall x \neq x_e, \tag{5.13}$$

such that

$$\dot{V}(\tilde{x}, u) \leqslant 0 \text{for all } x, u \text{ satisfying:} \tag{5.14}$$

$$V(\tilde{x}) \geq 1, (3.3), (5.11) \tag{5.15}$$

where $\tilde{Q} \geq 0$ is a symmetric matrix.

In the presence of a control input $u(t)$, the time derivative $\dot{V}(\tilde{x}, u)$ of the Lyapunov function $V(\tilde{x}, u)$ defined in (5.13) is not (4.41), but it becomes:

$$\dot{V}(\tilde{x}, u) = \begin{bmatrix} \tilde{x} \\ u \end{bmatrix}^\top \begin{bmatrix} \tilde{A}^\top \tilde{Q} + \tilde{Q}\tilde{A} & \tilde{Q}\tilde{B} \\ \tilde{B}^\top \tilde{Q} & 0_{m,m} \end{bmatrix} \begin{bmatrix} \tilde{x} \\ u \end{bmatrix}. \tag{5.16}$$

Notice that if $\dot{V}(\tilde{x}) \leqslant 0$ for all x, u satisfying $V(\tilde{x}) \geq 1$, then $V(\tilde{x}(t))$ defines a time-invariant ellipsoidal κ-level set from which the state trajectory will never go out over the infinite-time horizon; so if the set of initial conditions is within this level set, the state trajectory stemming from them will never exit this level set. In fact, to assure that the state trajectory $x(t)$ of the dynamical system in (3.3) does not leave the ellipsoid:

$$\mathcal{E}_x(Q, \bar{x}) = \{x \in \mathbb{R}^n | \quad (x - \bar{x})^\top Q(x - \bar{x}) \leq 1\} \tag{5.17}$$

it is sufficient to require the fulfillment of the inequality $\dot{V}(\tilde{x}(t)) \geq 0$, $\forall\, x(t)$, $u(t)$ satisfying $V(\tilde{x}(t)) \leq 1$, so that $V(\tilde{x}(t))$ decreases when $V(\tilde{x}(t))$ is outside the ellipsoid in (5.17). This condition can be rewritten as: $\dot{V}(\tilde{x}(t)) \leq 0$ for all x, u satisfying $V(\tilde{x}(t)) \geq 1$. It is obvious that this condition can be satisfied on the infinite-time horizon if and only if the dynamical system is Lyapunov stable. If not Lyapunov stable the state trajectory will not converge, as energy is added to the dynamical system from $u(t)$.

In addition to this, the state trajectory is constrained to remain in the service set $\mathcal{E}_s(\Xi, \bar{\xi})$, as aging attacks are launched while maintaining the desired service to guarantee the stealthiness of the attack. Therefore, the constraint in (5.12) is added to the requirement in (5.15), under the Assumption 1 that is specified for this method as follows:

Assumption 2. *There exists a control signal $u(t) \in \mathcal{E}_u(R, \bar{u})$ such that $x(t) \in \mathcal{E}_s(\Xi, \bar{\xi})$ for all $t \geq 0$ and for any initial conditions $x(t_0) \in \mathcal{E}_x(Q(t_0), \bar{x}(t_0))$.*

Finally, the problem we want to solve can be summarized as follows:

$$P_1 \begin{cases} \text{Find } \tilde{Q} \text{ such that} \\ \dot{V}(\tilde{x}, u) \leq 0 \quad \forall\, x(t),\, u(t) \text{ satisfying:} \\ V(\tilde{x}) \geq 1,\, u(t) \in \mathcal{E}_u(R, \bar{u}),\, x(t) \in \mathcal{E}_s(\Xi, \bar{\xi}) \\ \text{under Assumption 2} \\ \text{with } \tilde{Q} \geq 0 \end{cases} \tag{5.18}$$

Finally, Theorem 6 casts the problem P_1 into a SDP feasability problem with LMI constraints.

Theorem 6. *Consider an affine time-invariant continuous-time dynamical system according to (3.2) or equivalently (3.3), with matrices defined according to (3.4).*
If there exist $\tilde{Q} \in \mathbb{R}^{(n+1)\times(n+1)}$, *and* α, β, $\lambda \in \mathbb{R}$ *for which the following constraints are satisfied:*

$$-M - \alpha N - \beta S - \lambda T \geq 0, \tag{5.19}$$

$$\alpha, \beta, \lambda \geq 0 \tag{5.20}$$

with

$$M = \begin{bmatrix} \tilde{A}^\top \tilde{Q} + \tilde{Q}\tilde{A} & \tilde{Q}\tilde{B} \\ \tilde{B}^\top \tilde{Q} & 0_{m,m} \end{bmatrix} \tag{5.21}$$

$$N = \begin{bmatrix} \tilde{Q} - \Gamma^\top \Gamma & 0_{n+1,m} \\ 0_{m,n+1} & 0_{m,m} \end{bmatrix} \tag{5.22}$$

$$S = \begin{bmatrix} \Gamma^\top \Gamma - \Gamma^\top \bar{u}^\top R \bar{u} \Gamma & \Gamma^\top \bar{u}^\top R \\ R\bar{u}\Gamma & -R \end{bmatrix} \tag{5.23}$$

$$T = \begin{bmatrix} \Gamma^\top \Gamma - \tilde{\Xi} + \tilde{\Xi}\tilde{\xi}\Gamma + \Gamma^\top \tilde{\xi}^\top \tilde{\Xi} - \Gamma^\top \tilde{\xi}^\top \tilde{\Xi} \tilde{\xi} \Gamma & 0_{n+1,m} \\ 0_{m,n+1} & 0_{m,m} \end{bmatrix} \tag{5.24}$$

$$\tag{5.25}$$

and Γ, $\tilde{\Xi}$, $\tilde{\xi}$ *defined as in (5.6), then*

$$\tilde{x}(t_0)^\top \tilde{Q} \tilde{x}(t_0) \leqslant 1 \Rightarrow \tilde{x}(t)^\top \tilde{Q} \tilde{x}(t) \leqslant 1 \, \forall t \geq t_0, \tag{5.26}$$

under the constraints $\forall t \geq t_0$, $u(t) \in \mathcal{E}_u(R, \bar{u})$, $x(t) \in \mathcal{E}_s(\Xi, \xi)$.

Proof. Consider first (5.19); left and right multiply by $[\tilde{x}(t)^\top, u(t)^\top]^\top$, and consider αN, βS and λT as S-procedure terms by positive multiplier α, β and λ as in (5.20); this implies (Lemma 1):

$$[\tilde{x}^\top, u^\top] - M \, [\tilde{x}^\top, u^\top]^\top = -\dot{V}(\tilde{x}, u) \geqslant 0 \Leftrightarrow \dot{V}(\tilde{x}, u) \leqslant 0 \tag{5.27}$$

when

$$[\tilde{x}^\top, u^\top] \, N \, [\tilde{x}^\top, u^\top]^\top = V(\tilde{x}) - 1 \geqslant 0 \Leftrightarrow V(\tilde{x}) \geqslant 1. \tag{5.28}$$

$$[\tilde{x}^\top, u^\top] \, S \, [\tilde{x}^\top, u^\top]^\top \geqslant 0 \Leftrightarrow u(t) \in \mathcal{E}_u(R, \bar{u}). \tag{5.29}$$

$$[\tilde{x}^\top, u^\top] \, T \, [\tilde{x}^\top, u^\top]^\top \geqslant 0 \Leftrightarrow x(t) \in \mathcal{E}_x(\Xi, \bar{\xi}), \tag{5.30}$$

This means that the value of $V(\tilde{x})$ can only increase under the stated constraints, i.e. $V(\tilde{x}(t_0)) \leqslant 1 \Rightarrow V(\tilde{x}(t)) \leqslant 1 \; \forall t \geq 0$, which is the theorem statement. □

Due to the product of α with $\tilde{Q}$, two unknown variables, the matrix inequality in (5.45) in Theorem 6 is not a LMI, but a BMI. To deal with it, the constrained invariant state ellipsoid is computed iteratively for a fix $\alpha \geq 0$ at each iteration.

Ideally, we would like to design an iterative algorithm computing at each iteration the optimal α minimizing the size of $\mathcal{E}_x$, so that the iterative algorithm stops once the optimal α converges. At least, we would like to impose an upper bound $\bar{\alpha}$ on α to explore a finite interval.

Theorem 7. *Consider the matrix of the constraint (5.19) in Theorem 6 restated in (5.31)*

$$-M - \alpha N - \beta S - \lambda T \tag{5.31}$$

with α, β, $\gamma \geq 0$, and matrices M, N, S, and T are defined in (5.21), (5.22), (5.23), (5.24).

For the special case $\Xi = 0$, it is a necessary condition for the satisfaction of the constraint that

$$\alpha \leq -2\bar{\lambda}(\bar{A}) \tag{5.32}$$

where $\bar{\lambda}(\bar{A})$ denotes the largest eigenvalue of $\bar{A}$, with $\bar{A} = (A^\top + A)/2$, and A is the matrix of the dynamical system in (3.3)

Proof. Before starting with the proof, let us recall the following Lemmas.

Lemma 1. *If a matrix Y is positive semidefinite, the diagonal elements Y_{nn} are nonnegative.*

Lemma 2. *If a matrix Y is positive (negative) semidefinite, the trace is positive (negative).*

For the matrix in (5.31) to be positive semidefinite, it is necessary that the diagonal entries are nonnegative from Lemma 1. By expanding the vector $[\tilde{x}^\top, u^\top]$ into $[x^\top, 1, u^\top]$, matrices M, N, S, T in (5.19) can be rewritten:

$$M = \begin{bmatrix} A^\top Q + QA & \bullet & \bullet \\ \bullet & -a^\top Q\bar{x} - \bar{x}^\top Qa & \bullet \\ \bullet & \bullet & 0_{m,m} \end{bmatrix}$$

$$N = \begin{bmatrix} Q & \bullet & \bullet \\ \bullet & \bar{x}^\top Q\bar{x} - 1 & \bullet \\ \bullet & \bullet & 0_{m,m} \end{bmatrix}$$

$$S = \begin{bmatrix} 0_{n,n} & \bullet & \bullet \\ \bullet & 1 - \bar{u}^\top R\bar{u} & \bullet \\ \bullet & \bullet & -R \end{bmatrix}$$

$$T = \begin{bmatrix} -\Xi & \bullet & \bullet \\ \bullet & 1 - \bar{\xi}\Xi\bar{\xi} & \bullet \\ \bullet & \bullet & 0_{m,m} \end{bmatrix}$$

Thus, the necessary condition is:

$$-A^\top Q - QA - \alpha Q + \lambda \Xi \geq 0 \tag{5.33}$$
$$\Leftrightarrow \quad A^\top Q + QA + \alpha Q - \lambda \Xi \leq 0 \tag{5.34}$$

From Lemma 2, it yields:

$$\begin{aligned} & \operatorname{trace}(A^\top Q + QA + \alpha Q - \lambda \Xi) \leq 0 && (5.35)\\ \Leftrightarrow \quad & \operatorname{trace}(A^\top Q + QA) \leq \operatorname{trace}(-\alpha Q + \lambda \Xi) \\ \Leftrightarrow \quad & 2\operatorname{trace}(A^\top Q) \leq -\alpha \operatorname{trace}(Q) + \lambda \operatorname{trace}(\Xi) \end{aligned}$$

As $\lambda_n(\bar{A^\top})\operatorname{trace}(Q) \leq \operatorname{trace}(A^\top Q)$ where $\lambda_j(\bar{A})$ denotes the j^{th} largest eigenvalue of $\bar{A}$ with $\bar{A^\top} = (A^\top + A)/2$ then, (5.35) is equivalent to

$$2\lambda_n(\bar{A^\top})\operatorname{trace}(Q) \leq -\alpha \operatorname{trace}(Q) + \lambda \operatorname{trace}(\Xi)$$

In the special case $\Xi = 0$, it yields

$$2\lambda_n(\bar{A^\top}) \leq -\alpha \Leftrightarrow \alpha \leq -2\lambda_n(\bar{A^\top})$$

Then, it implies that $\alpha \leq \bar{\alpha}$ with $\bar{\alpha} = -2\lambda_n((A^\top + A)/2)$ for the special case $\Xi = 0$, which is the theorem statement. □

After having given the theorem providing sufficient conditions for a constrained invariant state ellipsoid in which the state remains over the infinite-time horizon, the three steps to analyze and prevent stealthy aging attacks are detailed, one for each subsection. Given a dynamical system as in (3.3), the input limitation $\mathcal{E}_u(R, \bar{u})$, the service set $\mathcal{E}_s(\Xi, \bar{\xi})$, and the abnormal aging set D_x the following steps are proposed.

5.3.3 First step: analysis of AT1

This first step aims to analyze stealthy aging attacks of AT1. To do so, we propose computing the constrained invariant state ellipsoid $\mathcal{E}_x(Q, \bar{x})$ for the given dynamical system in (3.3) and for a given control input set $\mathcal{E}_u(R, \bar{u})$. This constrained invariant state ellipsoid encompasses the constrained infinite time reachable set $\mathcal{X}(\infty, x(0))$ under the condition of existence of a control input $u(t) \in \mathcal{E}_u(R, \bar{u})$ that can constrain the state trajectory to remain inside the service set, $x(t) \in \mathcal{E}_s(\Xi, \bar{\xi})$ (Assumption 2).

As invariant sets are outer-approximation of the infinite-time reachable set $\mathcal{X}(\infty, x(0))$, there exist an infinite number of sets being invariant for the dynamical system. However, we are interested in the smallest constrained invariant state ellipsoid to have the best approximation, i.e. as close as possible to $\mathcal{X}(\infty, x(0))$. In fact, the lower the approximation error, the better the analysis of AT1 is. This can be obtained by maximizing the trace of $\tilde{Q}$ under Theorem 6 as stated in the convex Optimization Problem (OP) $\mathbf{OP_1}$.

$$\begin{aligned} \mathbf{OP_1}\text{:} \qquad & \underset{\tilde{Q},\beta,\lambda}{\text{maximize}} && \operatorname{trace}(\tilde{Q}) \\ & \text{subject to} && (5.19), (5.20) \end{aligned}$$

Note that maximizing the trace does not guarantee to find the tighest ellipsoid as the trace(.) operator relates to the size of the semiaxes. Hence, Algorithm 1 is proposed for finding the smallest constrained invariant state ellipsoid for the dynamical system in (3.3) with a control input set $\mathcal{E}_u(R, \bar{u})$ for a given α.

Algorithm 1 : $[Q, \bar{x}]$=Alg1(α, R, $\bar{u}$, Ξ, $\bar{\xi}$, $\mathbf{I_1}$)

Input: α, R, $\bar{u}$, Ξ, $\bar{\xi}$, $\mathbf{I_1}$:=$\{\tilde{A}, \tilde{B}\}$
Initial condition: -
2. Solve $\mathbf{OP_1}$
3. $[Q, \bar{x}]$=Normalize($\tilde{Q}$)
Output: Q, $\bar{x}$

Remark 9. *Notice that the numerical procedure will have a limited numerical precision, for this reason a normalization step (step 3.) has been introduced; typically the ellipsoids matrices might not fit into the form of (4.22), with a non-matching lower-right entry that corresponds to defining a set* $\{v \in \mathbb{R}^n \,|\, (v-\bar{v})^\top A^{nn}(v-\bar{v}) \leqslant \rho\}$ *with* $\rho \neq 1$. *The function Normalize*($\tilde{A}^{nn}$) *transforms a non-normalized (nn) ellipsoid defined as* $\mathcal{E}_v^{nn}(A^{nn}, \bar{v}) = \{v \in \mathbb{R}^n \,|\, (v-\bar{v})^\top A^{nn}(v-\bar{v}) \leqslant \rho\}$ *into a normalized one* $\mathcal{E}_v(A, \bar{v})$ *(*$\rho = 1$*).*

Thereafter, Algorithm 2 is proposed for finding the smallest constrained invariant state ellipsoid $\mathcal{E}_x$ for the dynamical system in (3.3) with a control input set $\mathcal{E}_u(R, \bar{u})$ over a value interval for α. Recall that we are interested in the smallest $\mathcal{E}_x$ to have the best approximation of X. The idea of Algorithm 2 is to recursively call Algorithm 1 to compute the smallest constrained invariant state ellipsoid $\mathcal{E}_x(Q, \bar{x})$ for a given α (step 1.1.). The size of the computed $\mathcal{E}_x(Q, \bar{x})$ is then compared with $\mathcal{E}_x(Q_{min}, \bar{x}_{min})$, that is the smallest ellipsoids found over the value interval of α. If the computed ellipsoid is smaller (step 1.2.), then update $\mathcal{E}_x(Q_{min}, \bar{x}_{min})$ with the recently computed ellipsoid (step 1.2.1.). The Algorithm repeats the same process for another value of α over the interval value. Note that Δ_α is the α step at each iteration such that α is updated at $\alpha + \Delta_\alpha$. After having explored the whole interval value of α, Algorithm 2 finally provides the smallest constrained invariant state ellipsoid $\mathcal{E}_x(Q_{min}, \bar{x}_{min})$ found at some α, if there exists. So, for the given dynamical system, $x(t_0) \in \mathcal{E}_x(Q_{min}, \bar{x}_{min})$ implies that $x(t) \in \mathcal{E}_x(Q_{min}, \bar{x}_{min})$, $\forall u(t) \in \mathcal{E}_u(R, \bar{u})$ and for $x(t) \in \mathcal{E}_s(\Xi, \bar{\xi})$, $\forall t \geq t_0$.

Remark 10. *As previously mentioned, the choice of* $\bar{\alpha}$ *for the general case, i.e.* $\Xi \neq 0$*, is still an open problem. Then, one can choose it arbitrarily.*

Algorithm 2 : $[Q_{min}, \bar{x}_{min}]$=Alg2(R, $\bar{u}$, Ξ, $\bar{\xi}$, $\mathbf{I_1}$, $\mathbf{I_2}$)

Input: $\underline{\alpha}$, $\bar{\alpha}$, Δ_α, R, $\bar{u}$, Ξ, $\bar{\xi}$, $\mathbf{I_2}$:={$\underline{\alpha}$, $\bar{\alpha}$, Δ_α}
Initial condition: $Q_{min} = 0_{n,n}$, $\bar{x}_{min} = 0_{n,1}$
1. For $\alpha = \underline{\alpha}$ to $\bar{\alpha}$ with a step Δ_α
  1.1. $[Q, \bar{x}]$=Alg1(α, R, $\bar{u}$, Ξ, $\bar{\xi}$, $\mathbf{I_1}$)
  1.2. If trace($\tilde{Q}$) > trace($\tilde{Q}_{min}$)
    1.2.1. Set $\tilde{Q}_{min} = \tilde{Q}$
Output: Q_{min}, $\bar{x}_{min}$

5.3.4 Second step: reachability analysis

This second step aims to check if the state trajectory can reach the abnormal aging set. This can be done by checking if the the constrained invariant state ellipsoid $\mathcal{E}_x(Q_{min}, \bar{x}_{min})$ resulting from the first step intersects with the abnormal aging set D_x. Instead of considering the intersection between both sets, we focus on the overlapping distance between an ellipsoid and an hyperplane defined in (4.27), as D_x is defined by the union of i number of hyperplane. Hence, we define the function Hit(Q_x,$\bar{x}$,$\mathcal{H}^i$) for determining if an ellipsoid $\mathcal{E}_v(A, \bar{v})$ hits the abnormal aging set $\mathcal{D}_x$. This function verifies that the distance from the ellipsoid to each hyperplane is negative. If it is, then the ellipsoid does not intersect with the abnormal aging set; otherwise it does.

Definition 34. *Function Hit: [hit]=Hit(Q_x, $\bar{x}$, $\mathcal{H}^i$)*

Input: Q_x, $\bar{x}$, $\mathcal{H}^i$
Initial condition: -
1. For $j = 1$ *to* i
  1.1. If dist$(\mathcal{E}_x(Q_x, \bar{x}), \mathcal{H}_j(c_j, \mu_j)) \geq 0$
    Set $hit = True$, *Goto Output*
  1.2. Else Set $hit = False$, *Goto Output*
Output: hit

Algorithm 3 implements this boolean function Hit(A,$\bar{v}$,$\mathcal{H}^i_v$) to check if the smallest constrained invariant state ellipsoid intersects with the abnormal aging set (step 2.). The output of the function and so the Algorithm 3 answers to RQ1 about the feasability of stealhty aging attacks of AT1.

Figure 5.5 is an illustration of RQ1 on a projection of two states and two inputs. The constrained invariant state ellipsoid $\mathcal{E}_x$ (dot pattern) is obtained for a given $\mathcal{E}_u$ by applying Algorithm 3. The intersection between $\mathcal{E}_x$ and the abnormal aging set D_x is studied (red polygon).

Algorithm 3 : $[hit]$=Alg3(R, $\bar{u}$, Ξ, $\bar{\xi}$, $\mathbf{I_1}$, $\mathbf{I_2}$, $\mathbf{I_3}$)

Input: R, $\bar{u}$, Ξ, $\bar{\xi}$, $\mathbf{I_1}$, $\mathbf{I_2}$, $\mathbf{I_3}$:=$\{\mathcal{H}_i, \forall i\}$
Initial condition: -
1. $[Q_{min}$,$\bar{x}_{min}]$ = Alg2(R, $\bar{u}$, Ξ, $\bar{\xi}$, $\mathbf{I_1}$, $\mathbf{I_2}$)
2. If Hit(Q_{min},$\bar{x}_{min}$,$\mathcal{H}_i$) = $True$
   2.1. Set $hit = True$, goto Output
3. Else Set $hit = False$, goto Output
Output: hit

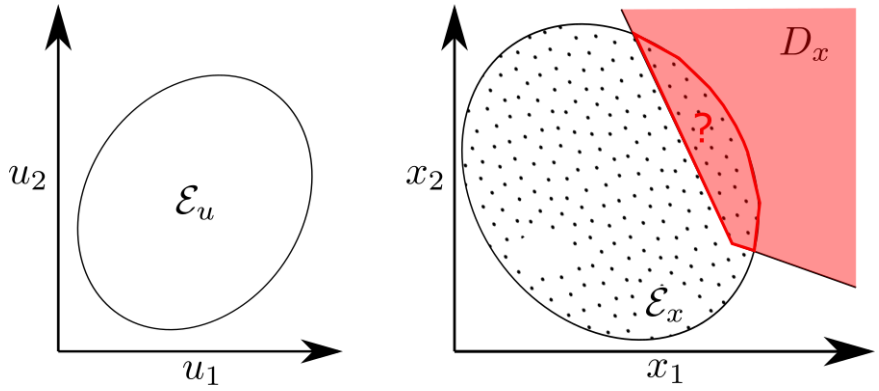


Figure 5.5: RQ1- Constrained invariant state set $\mathcal{E}_x$ (dot pattern) for the input set $\mathcal{E}_u$ by applying Algorithm 3, the abnormal aging set D_x (red area), and ($\mathcal{E}_x \cap D_x$) (red polygon)

5.3.5 Third step: restriction of the control input set

This third step aims to restrain the control input set $\mathcal{E}_u$ to $\hat{\mathcal{E}}_u$ such that the state trajectory cannot reach the abnormal aging set D_x for the dynamical system in (3.3) in the infinite-time horizon. In other words, the objective is to make AT1 infeasible, i.e. output of Algorithm 3 is false.

To do so, Algorithm 4 is proposed for finding $\hat{\mathcal{E}}_u$ such that the corresponding smallest constrained invariant state ellipsoid $\mathcal{E}_x(Q_{min}, \bar{x}_{min})$ (output of Algorithm 2) allows Hit(Q_{min},$\bar{x}_{min}$,$\mathcal{H}_i$) to result False.

Remark 11. *One could always choose $\hat{\mathcal{E}}_u(0, \bar{u}$, prohibiting any trajectory of the control signal $u(t)$. However, it also prohibits the control of the dynamical system, which is not desired as it is the purpose of the control system to deliver a service.*

Therefore, we want to find the largest control input set $\hat{\mathcal{E}}_u \subseteq \mathcal{E}_u$. Notice that we do not search for an optimal $\bar{u}$ and we let $\bar{u}$ be the same as for $\mathcal{E}_u$. By setting R as an additional decision variables of the optimization problem $\mathbf{OP_1}$, the constraint becomes a BMI due to the term βS, where S includes the shape matrix R. To tackle this issue, we propose to implement a bisection method on the scaling term δ for R for solving the quasi-convex problem. In fact, larger the control input set $\hat{\mathcal{E}}_u$, larger the smallest constrained invariant state ellipsoid is for a dynamical system as in (3.3). As an interpretation, larger

the control input set, larger the potential of energy acting on the system is; so the state trajectory can go further from the equilibrium point.

Hence, the main idea of Algorithm 4 is to perform the first and second steps for a fixed R. If the smallest constrained invariant state set does not intersect with the abnormal aging set, then increase the size of $\mathcal{E}_u$, i.e. R decreases; otherwise decrease, i.e. R increases. As an interpretation, by increasing the size of the control input we let the minimal constrained invariant state set to grow, whereas by decreasing its size we let the minimal constrained invariant state set to shrink.

Remark 12. *By fixing $\bar{u}$, it might occur in some cases that $\bar{u}$ is not optimal to find the largest $\hat{\mathcal{E}}_u$.*

The interval value of R is defined such that at the minimal size of $\hat{\mathcal{E}}_u$ the abnormal aging set is never hit by the smallest constrained invariant state ellipsoid, and at the maximal size $\hat{\mathcal{E}}_u = \mathcal{E}_u$ the abnormal aging set is hit by the constrained invariant state set. This allows the bisection method to find an optimal control input set. Note that from a physical viewpoint there always exists a feasible input set by letting R having a zero-size; otherwise it means that the abnormal aging set can always be reached.

Algorithm 4 converges, with respect to a tolerance tol, to the largest control input set $\hat{\mathcal{E}}_u(\delta R_0, \bar{u})$ for a sufficiently large R_0 such that $\mathcal{E}_x \cap D_x = \emptyset$, and a $\underline{\delta} \in [0, 1]$ sufficiently small such that $\mathcal{E}_x \cap D_x \neq \emptyset$ for $\mathcal{E}_u(\underline{\delta} R_0, \bar{u})$. Finally, for the given dynamical system, $x(t_0) \in \mathcal{E}_x(Q_{min}, \bar{x}_{min})$ implies that $x(t) \in \mathcal{E}_x(Q_{min}, \bar{x}_{min})$, $\forall u(t) \in \mathcal{E}_u(\delta R_0, \bar{u})$ and for $x(t) \in \mathcal{E}_s(\Xi, \bar{\xi})$, $\forall t \geq t_0$ and $\mathcal{E}_x \cap D_x = \emptyset$.

Algorithm 4 :[$diag$, δ_{wrk}]=Alg4(R_0,$\underline{\delta}$,tol,$\bar{u}$,Ξ,ξ, $\mathbf{I_1}$,$\mathbf{I_2}$, $\mathbf{I_3}$)

Input: R_0,$\underline{\delta}$,tol,$\bar{u}$,Ξ,ξ, $\mathbf{I_1}$,$\mathbf{I_2}$,$\mathbf{I_3}$
Initial condition: $diag = NotFound$
1. Set $\overline{\delta} = 1$
2. [hit]=Alg3($\underline{\delta} R_0$,$\bar{u}$,Ξ,ξ,$\mathbf{I_1}$,$\mathbf{I_2}$,$\mathbf{I_3}$)
3. If $hit = False$
   3.1. Set $diag = NeverHit$, $\delta = \underline{\delta}$, goto Output
4. [hit]=Alg3($\overline{\delta} R_0$,$\bar{u}$,Ξ,ξ,$\mathbf{I_1}$,$\mathbf{I_2}$,$\mathbf{I_3}$)
5. If $hit = True$
   5.1 Set $diag = AlwaysHit$, $\delta = \emptyset$, goto Output
6. Set $\delta = \overline{\delta}$
7. While $(\overline{\delta} - \underline{\delta} > tol)$
   7.1. $\delta_{it} = (\overline{\delta} + \underline{\delta})/2$
   7.2. $R = \delta_{it} R_0$
   7.3. [hit]=Alg3(R,$\bar{u}$,Ξ,ξ,$\mathbf{I_1}$,$\mathbf{I_2}$,$\mathbf{I_3}$)
   7.4. If $hit = True$
      7.4.1. Set $\underline{\delta} = \delta_{it}$
   7.5. Else Set $\overline{\delta} = \delta_{it}$, $diag = Found$, $\delta = \delta_{it}$
Output: $diag$, δ

Figure 5.6 is an illustration of RQ2 on a projection of two states and two inputs. The

current input limitation $\mathcal{E}_u$ and the smallest constrained invariant state set $\mathcal{E}_x$ (dot pattern) of RQ1 are drawn in solide lines. The largest restrained input set $\hat{\mathcal{E}}_u$ obtained from Algorithm 4 such that constrained invariant state set $\mathcal{E}_x$ (dot pattern) does not intersect with the abnormal aging set D_x.

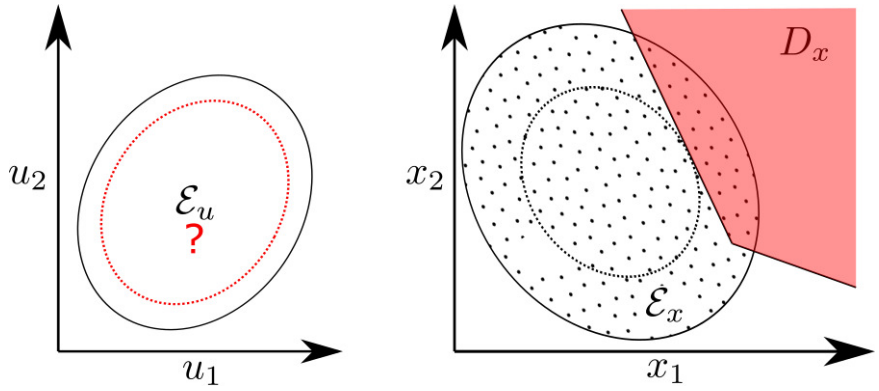


Figure 5.6: RQ2- Largest restrained input set $\hat{\mathcal{E}}_u$ with the corresponding constrained invariant state set $\mathcal{E}_x$ (dot pattern) in dotted lines by applying Algorithm 4, and RQ1 in solid lines, and the abnormal aging set D_x (red area)

One of the main limitation of the presented method is the necessary condition for the system to be Lyapunov stable. In many cases, the dynamical system is not Lyapunov stable without considering the closed-loop controller in the model in (3.3). The following Section presents a set-theoretic method that tackle this limitation.

5.4 Convex robust simulation-based method

To address the limitation of the previous method, we have proposed in [Escudero *et al.* 2020a] has proposed a robust convex simulation-based method which relies on computing the evolution of bounding state ellipsoids with respect to the time by means of LMI theory.

5.4.1 Robust simulation

By robust simulation we mean simulation of a dynamical system for a whole set of initial conditions, under a certain number of constraints. Namely, the problem of robust simulation consists in finding a bounding set for the state $x(t)$ at a final time t_f, given a set of possible initial values for it at an initial time t_0; this under the hypotheses of a given dynamical equation and in the presence of additional constraints of different kind. The bounding state sets considered in this thesis are ellipsoids, denoted $\mathcal{E}_x(t_f)$ for the bounding set for the state $x(t)$ at time t_f. Therefore, bounding state ellipsoids encapsulates the state trajectory of a dynamical system at the finite-time horizon from an initial state ellipsoid, i.e. $\mathcal{R}(t_f, x(0)) \subseteq \mathcal{E}_x(t_f)$, with $x(0) \in \mathcal{E}_x(t_0)$.

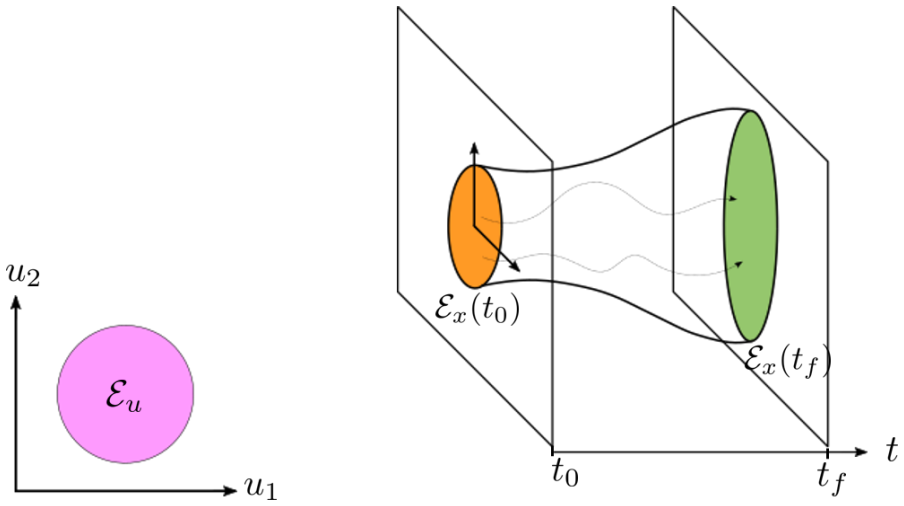


Figure 5.7: Bounding state ellipsoid $\mathcal{E}_x(t_f)$ from $\mathcal{E}_x(t_0)$

5.4.2 Main theorem of the convex robust simulation-based method

The main theorem of this method provides sufficient conditions for computing the constrained bounding state ellipsoid over a time interval $[t_0, t_f]$ for a dynamical system as in (3.3). This theorem will be subsequently employed in a set of algorithms, whose goal is the prediction and prevention of potential stealthy aging attacks of AT1.

This method is based on the search of a time-dependent pseudo-Lyapunov function [Ben-Talha *et al.* 2017]. The main idea is to define a positive definite function

$$V(\tilde{x}(t), t) = \tilde{x}(t)^\top \tilde{Q}(t)\tilde{x}(t) \geq 0 \ \forall x(t) \neq x_e, \text{ and } \forall t \in [t_0, t_f] \tag{5.36}$$

such that

$$\dot{V}(\tilde{x}(t), t) \leqslant 0, \text{for all } x(t), u(t), t \text{ satisfying:} \tag{5.37}$$

$$(3.3), (5.11), t \in [t_0, t_f] \tag{5.38}$$

and such that

$$V(\tilde{x}(t), t) \geq 0, \text{for all } t \text{ satisfying:} \tag{5.39}$$

$$t \in [t_0, t_f] \tag{5.40}$$

where $\tilde{Q}(t)$ is a time-varying symmetric matrix.

Time derivative of the pseudo-Lyapunov function $V(\tilde{x}(t), t)$ is given as follows:

$$\dot{V}(\tilde{x}(t), u(t), t) = \begin{bmatrix} \tilde{x}(t) \\ u(t) \end{bmatrix}^\top \begin{bmatrix} \dot{\tilde{Q}}(t) + \tilde{A}^\top \tilde{Q}(t) + \tilde{Q}(t)\tilde{A} & \tilde{Q}(t)\tilde{B} \\ \tilde{B}^\top \tilde{Q}(t) & 0_{m,m} \end{bmatrix} \begin{bmatrix} \tilde{x}(t) \\ u(t) \end{bmatrix} \tag{5.41}$$

Notice that if $\dot{V}(x(t), t) \leqslant 0$, then $V(\tilde{x}(t), t)$ defines time-variant ellipsoidal level sets from which the state trajectory will never go out over the given time interval; so if a set of initial conditions is within a level set, the trajectory stemming from them will never exit the level set of the same value. The time-variant $\tilde{Q}(t)$ allows the ellipsoidal level

sets to move and grow or shrink over time with the state trajectory, whether the state trajectory is converging or not.

In order to make the search tractable, we are going to limit it to the set of matrix-valued polynomials, i.e. we will set $\tilde{Q}(t) \in \mathbb{R}_{2d}^{(n+1)\times(n+1)}[t]$ for an arbitrarily chosen integer $d \geqslant 1$. It is clear that specifying $\tilde{Q}(t)$ in this way adds some conservatism, which can be progressively reduced by increasing the degree of the matrix-valued polynomial. $\dot{\tilde{Q}}(t)$ is then its first time-derivative, with $\dot{\tilde{Q}} \in \mathbb{R}_{2d-1}^{(n+1)\times(n+1)}[t]$.

As for the invariant-set based method, the state trajectory is constrained to remain in the service set by adding the constraint (5.12), under Assumption 1 that is specified for this method as follows:

Assumption 3. *There exists a control input signal $u(t) \in \mathcal{E}_u(R, \bar{u})$ such that $x(t) \in \mathcal{E}_s(\Xi, \bar{\xi})$ for all $t \in [t_0, t_f]$ and for initial conditions $x(t_0) \in \mathcal{E}_x(Q(t_0), \bar{x}(t_0))$.*

If the assumption above is not satisfied, it means that no aging attack can be performed without being detected, which put us on the safe side.
Subsequently, the constraint of $t \in [t_0, t_f]$ can be expressed as follows

$$\gamma(t) = (t_f - t)(t - t_0) \geqslant 0 \tag{5.42}$$

Finally, the problem we want to solve can be summarized as follows:

$$P_2 \begin{cases} \text{Find } \tilde{Q}(t) \text{ such that} \\ (i) V(\tilde{x}(t), t) \geq 0 \quad \text{for } t \in [t_0, t_f], \\ (ii) \dot{V}(\tilde{x}(t), u(t), t) \leq 0 \quad \forall x(t),\, u(t) \text{satisfying:} \\ u(t) \in \mathcal{E}_u(R, \bar{u}),\, x(t) \in \mathcal{E}_s(\Xi, \bar{\xi}),\, t \in [t_0, t_f] \\ \text{under Assumption 1} \\ \text{with } \tilde{Q} \geq 0 \end{cases} \tag{5.43}$$

Finally, Theorem 8 casts the problem P_2 into a SDP feasability problem with LMI constraints. Before formulating it, consider

$$\tilde{I} = \begin{bmatrix} I_n & 0_{n,1} \\ 0_{1,n} & 0 \end{bmatrix}, \tilde{\Xi} = \begin{bmatrix} \Xi & 0_{n,1} \\ 0_{1,n} & 0, \end{bmatrix}, \tilde{\xi} = \begin{bmatrix} \xi \\ 0 \end{bmatrix}. \tag{5.44}$$

The main theorem is then stated as follows.

Theorem 8. *Consider an affine time-invariant continuous-time system according to (3.2) or equivalently (3.3), with matrices defined according to (3.4).*

If for a given $d \geqslant 1$ there exist $\tilde{Q}(t) \in \mathbb{R}_{2d}^{(n+1)\times(n+1)}[t]$, $Z(t), X(t) \in \mathbb{R}_{2d-2}^{(n+1)\times(n+1)}[t]$ $\alpha(t), \beta(t) \in \mathbb{R}_{2d-2}[t]$, and a scalar $\varepsilon > 0$ for which the following constraints are satisfied:

$$-W(t) - \alpha(t)S - \beta(t)T - \gamma(t)Z(t) \text{ is MSOS}, \tag{5.45}$$

$$\tilde{Q}(t) - \gamma(t)X(t) - \varepsilon\tilde{I} \text{ is MSOS} \tag{5.46}$$

$$Z(t),\, X(t) \text{ are MSOS} \tag{5.47}$$

$$\alpha(t), \beta(t) \text{ are SOS} \tag{5.48}$$

with

$$W(t) = \begin{bmatrix} \dot{\tilde{Q}}(t) + \tilde{A}^\top \tilde{Q}(t) + \tilde{Q}(t)\tilde{A} & \tilde{Q}(t)\tilde{B} \\ \tilde{B}^\top \tilde{Q}(t) & 0_{m,m} \end{bmatrix}$$
$$S = \begin{bmatrix} \Gamma^\top \Gamma - \Gamma^\top \bar{u}^\top R \bar{u} \Gamma & \Gamma^\top \bar{u}^\top R \\ R\bar{u}\Gamma & -R \end{bmatrix}$$
$$T = \begin{bmatrix} \Gamma^\top \Gamma - \tilde{\Xi} + \tilde{\Xi}\tilde{\xi}\Gamma + \Gamma^\top \tilde{\xi}^\top \tilde{\Xi} - \Gamma^\top \tilde{\xi}^\top \tilde{\Xi} \tilde{\xi} \Gamma & 0_{n+1,m} \\ 0_{m,n+1} & 0_{m,m} \end{bmatrix}$$

and Γ, Ξ, $\tilde{\xi}$ *defined as in* (5.6), *and* $\gamma(t)$, $\tilde{I}$ *defined as in* (5.42), (5.44), *then*

$$\tilde{x}(t_0)^\top \tilde{Q}(t_0)\tilde{x}(t_0) \leqslant 1 \Rightarrow \tilde{x}(t)^\top \tilde{Q}(t)\tilde{x}(t) \leqslant 1 \, \forall t \in [t_0, t_f], \tag{5.49}$$

under the constraints $\forall t \in [t_0, t_f] u(t) \in \mathcal{E}_u(R, \bar{u})$, $x(t) \in \mathcal{E}_x(\Xi, \xi)$.

Proof. Consider first (5.46); left and right multiply by $\tilde{x}(t)$, and consider $\gamma(t)X(t)$ as an S-procedure term by positive multiplier $X(t)$ as in (5.47); this implies (Lemma 1):

$$\tilde{x}(t)^\top \tilde{Q}(t)\tilde{x}(t) \geqslant \varepsilon ||x(t)||^2 \tag{5.50}$$

when $\gamma(t) \geqslant 0$, i.e. $V(x(t), t) = \tilde{x}(t)^\top \tilde{Q}(t)\tilde{x}(t) > 0$ when $x(t) \neq 0$, $t \in [t_0, t_f]$. Subsequently, consider (5.45); left and right multiply by $[\tilde{x}(t)^\top, u(t)^\top]^\top$, and consider $\alpha(t)S$, $\beta(t)T$ and $\gamma(t)Z(t)$ as S-procedure terms by positive multipliers $\alpha(t)$, $\beta(t)$ and $Z(t)$ as in (5.47) and (5.48); this implies (Lemma 1):

$$[\tilde{x}(t)^\top, u(t)^\top] \, W(t) \, [\tilde{x}(t)^\top, u(t)^\top]^\top = \dot{V}(\tilde{x}(t), t) \leqslant 0 \tag{5.51}$$

when

$$[\tilde{x}(t)^\top, u(t)^\top] \, S \, [\tilde{x}(t)^\top, u(t)^\top]^\top \geqslant 0 \Leftrightarrow u(t) \in \mathcal{E}_u(R, \bar{u}). \tag{5.52}$$
$$[\tilde{x}(t)^\top, u(t)^\top] \, T \, [\tilde{x}(t)^\top, u(t)^\top]^\top \geqslant 0 \Leftrightarrow x(t) \in \mathcal{E}_x(\Xi, \xi), \tag{5.53}$$

and when $\gamma(t) \geqslant 0$, i.e. when $t \in [t_0, t_f]$. This means that the value of $V(\tilde{x}(t), t)$ can only increase under the stated constraints, i.e. $V(\tilde{x}(t_0), t_0) \leqslant 1 \Rightarrow V(\tilde{x}(t), t) \leqslant 1$ in the considered interval, which is the theorem statement. □

After having given the theorem providing sufficient conditions for a bounding state ellipsoid at time t_f in which the state trajectory remains at t_f, the three steps to analyze and prevent stealthy aging attacks are detailed, one for each subsection. Given a dynamical system as in (3.3), the input limitation $\mathcal{E}_u(R, \bar{u})$, the service set $\mathcal{E}_s(\Xi, \bar{\xi})$, and the abnormal aging set D_x the following steps are proposed. As both methods are based on the same general idea described in Section 5.2.1, the algorithms we are going to present are rather similar to the ones for the invariant set based method. Thus, we will only present here the main differences.

In this section, we are proposing algorithms involving Theorem 8 for analyzing and preventing stealthy aging attacks over a given time interval. In other words, we want to restrain the input set $\mathcal{E}_u(R, \bar{u})$ into the set $\hat{\mathcal{E}}_u(\hat{R}, \bar{u})$, $(\mathcal{E}_u(\hat{R}, \bar{u}) \subseteq \mathcal{E}_u(R, \bar{u}))$ such that a dangerous set $\mathcal{D}_x$ is avoided over a time interval $[t_0, t_f]$.

5.4.3 First step: analysis of AT1

This first step aims to quantify the effect on the state trajectory $x(t)$ induced by AT1 over a time interval $[t_0, t_f]$. To do so, we propose to compute the constrained bounding state ellipsoid $\mathcal{E}_x(t_f)$ for the given dynamical system in (3.3) and for a given control input set $\mathcal{E}_u(R, \bar{u})$. This constrained bounding state ellipsoid encompasses the constrained finite-time reachable set $X(t_f, x(0))$ under the condition of existence of a control input $u(t) \in \mathcal{E}_u(R, \bar{u})$ that can constrain the state trajectory to remain inside the service set $\mathcal{E}_s(\Xi, \bar{\xi})$ over the time interval $[t_0, t_f]$, i.e. $x(t) \in \mathcal{E}_s(\Xi, \bar{\xi}) \, \forall t \in [t_0, t_f]$, (Assumption 3).

There exists many constrained bounding state ellipsoids from Theorem 8, i.e. a constrained bounding state ellipsoid is an outer-approximation of the constrained finite-time reachable set. Thus, we want to find the smallest bounding state ellipsoid to have the best approximation. As previously, this can be obtained by maximizing the trace of $\tilde{Q}(t_f)$ under Theorem 8 with initial conditions $(Q(t_0),\bar{x}(t_0))$ as stated in the convex optimization problem in $\mathbf{OP_2}$. So, for the given dynamical system, if $x(t_0) \in \mathcal{E}_x(t_0)$ defined with $(Q(t_0), \bar{x}(t_0))$, then it implies that $x(t) \in \mathcal{E}_x(t_f)$ defined with $(Q(t_f), \bar{x}(t_f))$, $\forall t \in [t_0, t_f]$, $\forall u(t) \in \mathcal{E}_u(R, \bar{u})$, and for $x(t) \in \mathcal{E}_s(\Xi, \bar{\xi})$.

In order to give more degrees of freedom to the search of such ellipsoids and without loss of generality, let t_0 be time 0 and split the final time t_f in $N \in \mathbb{N}$ time steps such that $N = t_f \Delta_t^{-1}$ with Δ_t a small time step, which simplifies the computations. Hence, Algorithm 1 is proposed for finding the minimal bounding ellipsoid for system in (3.3) with a control input set $\mathcal{E}_u(R, \bar{u})$) over the time interval $[0, \Delta_t]$ by using Theorem 8. Notice that (step 3.) is the normalization step due to limited numerical precision as detailed previously in Remark 9.

$$\mathbf{OP_2}: \quad \underset{\tilde{Q}_{\Delta_t}, Z(t), X(t), \alpha(t), \beta(t)}{\text{maximize}} \quad \text{trace}(\tilde{Q}_{\Delta_t})$$
$$\text{subject to} \quad (5.45), (5.46), (5.47), (5.48)$$

Algorithm 5 : $[Q_{\Delta_t}, \bar{x}_{\Delta_t}]$=Alg5$(Q(t_0),\bar{x}(t_0),R,\bar{u},\Xi,\xi, \mathbf{I_1})$

Input: $Q(t_0),\bar{x}(t_0),R,\bar{u}, \Xi, \xi, \mathbf{I_1}:=\{A, B, \Delta_t, d\}$
Initial condition: $Q_0 = Q(t_0)$, $\bar{x}_0 = \bar{x}(t_0)$
1. Set $\tilde{Q}_0 = \begin{bmatrix} Q_0 & -Q_0\bar{x}_0 \\ -\bar{x}_0^\top Q_0 & \bar{x}_0^\top Q_0 \bar{x}_0 \end{bmatrix}$,
$\tilde{Q}(t) = \sum_{i=0}^{2d} \tilde{Q}_i t^i$
2. Solve $\mathbf{P_1}$
3. $[Q(\Delta_t), \bar{x}(\Delta_t)]$=Normalize$(\tilde{Q}(\Delta_t))$
Output: $Q(\Delta_t)$, $\bar{x}(\Delta_t)$

5.4.4 Second step: reachability assessing

This second step aims to check if the state trajectory can reach the abnormal aging set D_x over the time interval $[t_0, t_f]$. Algorithm 5.4.4, whose goal is to determine if and at which time $t_\cap$ (with $t_\cap \in \{0, \Delta_t, ..., N\Delta_t\}$) the bounding ellipsoid hits the abnormal

aging set $\mathcal{D}_x$. Notice that $t_\cap$ could be exploited for further defensive strategies. The idea of Algorithm 5.4.4 is to compute the evolution of the smallest bounding state ellipsoids over the time interval $[t_0, t_f]$, and check if none of them intersects with D_x. To do so, Algorithm 2 implements an iterative procedure with N iterations that performs two steps at each iterations i:

1. The computation of the final (smallest) bounding state ellipsoid at time $t_0 + i\Delta_t$, $\mathcal{E}_x(t_0 + i\Delta_t)$ defined for $(Q(t_0 + i\Delta_t), \bar{x}(t_0 + i\Delta_t))$

2. The computation of the distance between $\mathcal{E}_x(t_0 + i\Delta_t)$ and D_x from the boolean function Hit(.) already presented previously in Section 5.3.4.

Finally, Algorithm 5.4.4 provides two arrays *hit* and $t_\cap$ to know at which time the smallest bounding state ellipsoid reaches the abnormal aging set. If all the element of *hit* are set at *False*, it means that stealthy aging attacks of AT1 are not feasible over the time interval $[t_0, t_f]$, otherwise they are (RQ1). In that last case, the third step is proposed to make AT1 not feasible.

Algorithm 6 : $[hit, t_\cap]$=Alg6$(Q(t_0), \bar{x}(t_0), R, \bar{u}, \Xi, \xi, \mathbf{I_1}, \mathbf{I_2})$

Input: $Q(t_0), \bar{x}(t_0), R, \bar{u}, \Xi, \xi, \mathbf{I_1}, \mathbf{I_2} := \{\mathcal{H}_x^i, t_f\}$
Initial condition: $Q_0 = Q(t_0)$, $\bar{x}_0 = \bar{x}(t_0)$
1. Set $N = t_f \Delta_t^{-1}$
2. For $j = 1$ to N
   2.1. $[Q_{\Delta_t}, \bar{x}_{\Delta_t}]$=Alg5$(Q_0, \bar{x}_0, R, \bar{u}, \Xi, \xi, \mathbf{I_1}, \mathbf{I_2})$
   2.2. If Hit$(Q_{\Delta_t}, \bar{x}_{\Delta_t}, \mathcal{H}_x^i) = True$
   Set $hit = True$, $t_\cap = j\Delta_t$, Goto Output
   2.3. Set $Q_0 = Q_{\Delta_t}$, $\bar{x}_0 = \bar{x}_{\Delta_t}$
3. Set $hit = False$, $t_\cap = \emptyset$
Output: hit, $t_\cap$

Figure 5.8 is an illustration of RQ1 on a projection of two states and two inputs. The evolution of bounding state ellipsoids over the time interval $[t_0, t_f]$ (dot pattern) are obtained for a given control input set $\mathcal{E}_u$ by applying Algorithm 6 with $N = 2$. At each iteration i, the intersection between the i^{th} bounding state ellipsoid and the abnormal aging set D_x is studied to determine the feasibility of stealthy aging attacks.

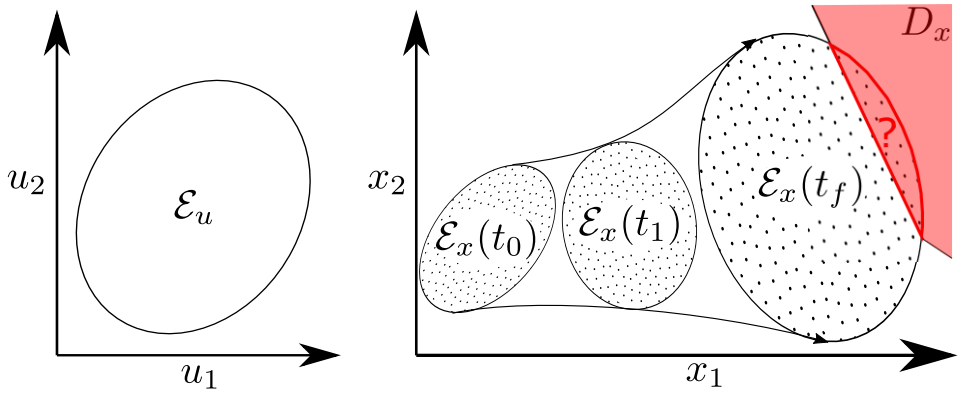


Figure 5.8: RQ1- Evolution of bounding state ellipsoids (dot pattern) over the time interval $[t_0, t_f]$ for the input set $\mathcal{E}_u$, the abnormal aging set D_x (red area), and $(\mathcal{E}_x(t) \cap D_x)$ (red polygon) by applying Algorithm 6

5.4.5 Third step: Restriction of the control input set

This third step aims to restrain the control input set $\mathcal{E}_u$ to $\hat{\mathcal{E}}_u$ such that the state trajectory cannot reach the abnormal aging set D_x for the dynamical system in (3.3) over the time interval $[t_0, t_f]$. In other words, the objective is to make AT1 unfeasible over the given time interval, i.e. output of Algorithm 5.4.4 is false for every element of *hit*.

For the same reason stated previously, the objective is to find the largest control input set $\hat{\mathcal{E}}_u$ such that the evolution of bounding state ellipsoids over the time interval $[t_0, t_f]$ never hits the abnormal aging set D_x. By setting R as an additional decision variables of the optimization problem $\mathbf{OP_2}$ becomes a BMI due to the term $\alpha(t)S$, where S includes the shape matrix R. Similarly as previously, we propose to implement a bisection method on the scaling term δ for R for solving the quasi-convex problem. Further details have already been given previously in Section 5.3.5. Hence, Algorithm 5.4.5 is proposed. It converges with respect to a tolerance *tol*, to the largest control input set $\mathcal{E}_u(\delta R_0, \bar{u})$ for a sufficiently large R_0 such that $\mathcal{E}_x(t_0 + i\Delta_t) \cap D_x = \emptyset$, $\forall i = 1, ..., N$, and a $\underline{\delta} \in [0, 1]$ sufficiently small such that $\mathcal{E}_x(t_0 + i\Delta_t) \cap D_x \neq \emptyset$ for $\mathcal{E}_u(\underline{\delta} R_0, \bar{u})$ for at least an iteration i.

Algorithm 7 :$[diag, \delta_{wrk}]$=
Alg7(R_0,$\underline{\delta}$,tol,$Q(t_0)$,$\bar{x}(t_0)$,$\bar{u}$,Ξ,ξ, $\mathbf{I_1}$,$\mathbf{I_2}$)

Input: R_0,$\underline{\delta}$,tol,$Q(t_0)$,$\bar{x}(t_0)$,$\bar{u}$,Ξ,ξ, $\mathbf{I_1}$,$\mathbf{I_2}$
Initial condition: $diag = NotFound$, $Q_0 = Q(t_0)$, $\bar{x}_0 = \bar{x}(t_0)$
1. Set $\overline{\delta} = 1$
2. $[hit, t_\cap]$=Alg6($Q(t_0)$,$\bar{x}(t_0)$,$\underline{\delta}R_0$,$\bar{u}$,Ξ,ξ,$\mathbf{I_1}$,$\mathbf{I_2}$)
3. If $hit = False$
   3.1. Set $diag = NeverHit$, $\delta = \underline{\delta}$, Goto Output
4. $[hit, t_\cap]$=Alg6($Q(t_0)$,$\bar{x}(t_0)$,$\overline{\delta}R_0$,$\bar{u}$,Ξ,ξ,$\mathbf{I_1}$,$\mathbf{I_2}$)
5. If $hit = True$
   5.1 Set $diag = AlwaysHit$, $\delta = \emptyset$, Goto Output
6. Set $\delta = \overline{\delta}$
7. While $(\overline{\delta} - \underline{\delta} > tol)$
   7.1. $\delta_{it} = (\overline{\delta} + \underline{\delta})/2$
   7.2. $R = \delta_{it} R_0$
   7.3. $[hit, t_\cap]$=Alg6($Q(t_0)$,$\bar{x}(t_0)$,R,$\bar{u}$,Ξ,ξ,$\mathbf{I_1}$,$\mathbf{I_2}$)
   7.4. If $hit = True$, Set $\underline{\delta} = \delta_{it}$
   7.4. Else, Set $\overline{\delta} = \delta_{it}$, $diag = Found$, $\delta = \delta_{it}$
Output: $diag$, δ

Figure 5.9 is an illustration of RQ2 on a projection of two states and two inputs. The evolution of bounding state ellipsoids over the time interval $[t_0, t_f]$ (dot pattern) for the current input limitation $\mathcal{E}_u$ of RQ1 are drawn in solid lines. The restrained input set $\hat{\mathcal{E}}_u$ is computed by applying Algorithm 7 with $N = 2$ such that the evolution of bounding state ellipsoids (dotted lines) do not hit with the abnormal aging set D_x.

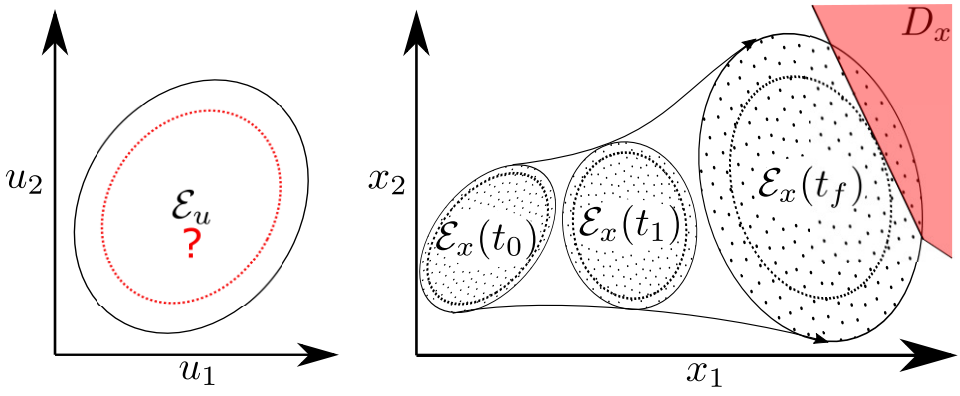


Figure 5.9: RQ2- Restrained input set $\hat{\mathcal{E}}_u$ with the corresponding evolution of bounding state ellipsoids (dot pattern) over the time interval $[t_0, t_f]$ in dotted lines by applying Algorithm 7, RQ1 in solid lines, and the abnormal aging set D_x (red area)

5.5 Conclusion

In this Chapter, we have proposed two set theoretic methods based on an outer-approximation of the reachable state set to answer RQ1 and RQ2 for stealthy aging attacks of AT1. The first one based on invariant set theory is limited to Lyapunov stable system whereas the last one based on robust convex simulation can be applied at any systems, i.e. Lyapunov stable or not. In the next Chapter, we will propose a method for analyzing stealthy aging attacks of a more general case (AT2).

Chapter 6

Analysis: general case of stealthy aging attacks

Contents

6.1 Introduction

In Chapter 5, we have detailed the methods for analyzing and preventing stealthy aging attacks of AT1. However, AT1 is a subcase of stealthy aging attacks. The general case is an attack that maximizes the time integral of a degradation function, instead of reaching an abnormal aging set. Hence, this Chapter proposes a method for analyzing stealthy aging attacks of AT2.

Section 6.2 formally defines the general case of stealthy aging attacks (AT1) to provide the framework of the problem the method will solve. Thereafter, the method for analyzing stealthy aging attacks (AT1) is detailed in Section 6.3.

6.2 Research question

From the general Definition 5 of stealthy aging attacks, our investigation concerns the following Attack Type.

Definition 35. ***(Attack Type 2 (AT2)).*** *The input $u(t)$ is manipulated by an attacker to maximize the time integral of a degradation function $g(x(t), u(t))$ while keeping the state trajectory inside the service set X_s. The degradation function models the stresses acting on the process.*

After having defined the stealthy aging attacks of AT2 we consider in this Chapter, the problem we address is explained in the following.

Consider a discretized version of (3.3), namely:

$$x_{k+1} = A_d x_k + B_d u_k + a_d \tag{6.1}$$

where u_k is a piecewise constant input, x_k is a piecewise constant state; and $A_d \in \mathbb{R}^{n\times n}$, $B_d \in \mathbb{R}^{n\times m}$, $a_d \in \mathbb{R}^n$
Lastly, consider $g(x_k, u_k)$ the discretized version of the degradation function $g(x(t), u(t))$. From this new type of attack, the following Research Question is answered in this Chapter:

Definition 36. ***(RQ3).*** *How the control signal u_k can be chosen to maximize the time integral of the degradation function $g(x_k, u_k)$, which we call cost degradation function, while constraining the state x_k to remain in the service set X_s.*

This research question differs from the previous method answering RQ1. In fact, this method assesses what kind of control signal u_k an attacker could replace to launch a stealthy aging attack. This is an assessment method instead of a faisability method as the previous ones.

For the same reason already presented in the previous Chapter, the input u_k is assumed bounded at every time instant due to physical input limitation:

$$u_k \in \mathcal{U}, \quad k \geq 0 \tag{6.2}$$

In addition to this, the state x_k is constrained by the service set X_s to remain stealthy:

$$x_k \in \mathcal{X}_s, \quad k \geq 0 \tag{6.3}$$

Remark 13. *Stealthy aging attacks of AT1 are a special case of AT2: the objective is to reach the abnormal aging set at some time instead of maximizing a cost degradation function. It can be addressed by setting a piecewise function g with non zero positive values only in the abnormal aging set D_x, and zero elsewhere.*

The next section presents the method for analyzing stealthy aging attacks of AT2, which includes AT1 as a subcase.

6.3 Analysis of stealthy aging attacks of AT2

6.3.1 Proposed method

The analysis can be seen as a design method where the optimal control signal is searched for. Then, the optimal stealthy aging attack is formulated as an optimization problem maximizing the time integral of the degradation function (6.4) while keeping the states inside the stealthy set X_s during the attack period of M time instants. The idea is to guarantee that the error of the delivered service with respect to the desired one is bounded by a given threshold δ sufficiently low such that it is not perceptible from the service state (e.g. measurements) (6.7), so that no detection occurs in the detector. Also, the input signal is constrained by the input limitations (6.6). After having launched the

attack, we want the dynamical system to return to its normal state in the attack-free case. To do so, we impose initial and final states to be equal to the normal ones.

OP_3: **Optimal abnormal input signal**

$$\underset{u_k, x_{k+1}}{\text{maximize}} \quad \sum_{k=0}^{M-1} g(x_k, u_k) \tag{6.4}$$

$$\text{subject to} \quad x_{k+1} = A_d x_k + B_d u_k + a_d \tag{6.5}$$

$$\underline{u}_k \leq u_k \leq \overline{u}_k \tag{6.6}$$

$$x_{k+1}^{des} - \delta \leq x_{k+1} \leq x_{k+1}^{des} + \delta \tag{6.7}$$

$$x_0 = x_0^{des}, x_N = x_N^{des} \tag{6.8}$$

6.3.2 Sequential Linear Programming

One issue of the previous optimization problems is the non-linearity of the degradation function $g(x_k, u_k)$. Indeed, aging of actuators involves most of the time a magnitude of a state (e.g. torque magnitude, current magnitude) which increases the stresses of the actuator.

To deal with the typical non-linearity of the degradation function g, we have proposed to solve the optimization problem $\mathbf{OP_3}$ with Sequential or Successive Linear Programming (SLP). SLP is an iterative procedure that approaches the solution of the original non-linear problem by generating a sequence of solutions of linear subproblems approximating it (LP problem) [Palacios-Gomez *et al.* 1982].

The main idea is to approximate the non-linear function $g(x_k, u_k)$ of the optimization problem by its Taylor series linearization at the operating points denoted by ^{0} [Vukic 2003]. For the sake of clarity, we drop out the k-instant notation; at each SLP iteration $g(x, u)$ is replaced by:

$$g(x^0, u^0) + \sum_{j=1}^{n} \frac{\partial g(x,u)}{\partial x_j}^0 \Delta x_j + \sum_{j=1}^{m} \frac{\partial g(x,u)}{\partial u_j}^0 \Delta u_j \tag{6.9}$$

where $\Delta x_j = x - x^0$ and $\Delta u = u - u^0$.

This linearization is only accurate for Δx_j and Δu_j sufficiently small; so upper and lower bounds are imposed as additional constraints in the optimization problem.

Hence, the optimization problems are transformed into a succession of linear programs in which the operating point x^0 and u^0 at the iteration j are respectively equals to the optimal $x^\star$ and $u^\star$ obtained at the iteration $j-1$. The SLP stops once Δx and Δu converges.

6.4 Conclusion

In this Chapter, we have proposed a method based on Linear Programming to answer RQ3 for stealthy aging attacks of AT2, which include AT1 as a subcase. In the next Part, we will apply the proposed methods for analyzing and preventing stealthy aging attacks of AT1 and AT2 upon two application After having detailed the methods for analyzing

and preventing stealthy aging attacks, the next Chapter proposes to apply them upon two application examples.

Part III

Simulation

Chapter 7

Application examples

Contents

7.1 Introduction

In this Chapter, the proposed methods for analyzing and preventing stealthy aging attacks of AT1 and AT2 are applied upon two application examples: a thin film manufacturing control and a satellite reorientation maneuver control. As we will see in this Chapter, both applications are relevant to study the proposed methods. In fact, the model of the thin film manufacturing system is Lyapunov stable, so that both methods can be applied and compared, whereas the satellite system is not Lyapunov stable and is over actuated. Besides, both applications exhibit a different system property: the satellite system is an over-actuated system, unlike the other application.

7.2 Thin film manufacturing

7.2.1 System description

Consider the film mfg. system represented in Figure 7.1. It represents the manufacture of continuous webs such as paper products, plastic films, or metal foils. In such manufacture, the objective is to move the continuous web in each part of the film mfg. system to perform some operations. The control objective is to maintain the tension and the

speed of the web at the reference values between the consecutive pairs of drive rolls. Film mfg. system typically consists of drive rollers and idler rollers. In a nutshell, they are cylindrically-shaped material handling components for conveying items through a machine, process, or environment. Drive rollers rotate via direct interaction with a motor, whereas idler rollers do not received direct mechanical input from a motor [Globalspec].

First and foremost, we detail briefly the physical relations between the variables to get the model of the system. For further details about the modeling, the interested reader can consult the reference [Chen *et al.* 2004]. As shown in Figure 7.1, the system is equipped with two drive rollers rx and ry having the same mechanical characteristics (B, J, R, D, K), detailed later. They are controlled respectively by two electric motors mx and my having the same electrical characteristics (R_a, L_a, K_e, K_t), detailed later. The torque produced by each motor is transmitted via a belt to the drive rollers. The drive roller torque is denoted $\tau_x(t)$, $\tau_y(t)$ for rx , ry, respectively. By considering similar motor characteristics and roller characteristics it allows to simplify the modeling (detailed for the general case in [Chen *et al.* 2004]). Note that in the following the notation m refers to a motor, and r to a roller.

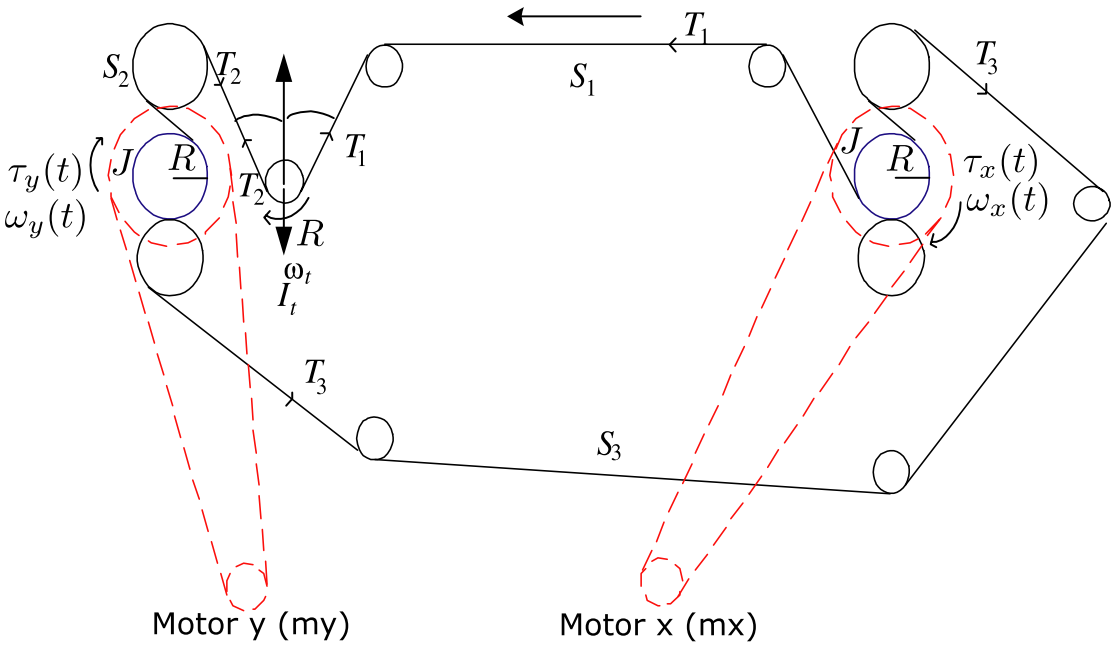


Figure 7.1: Film mfg. system (simplified from [Chen *et al.* 2004])

7.2.1.1 Web tension $T(t)$

In this subsection, we express the relation between the web tension $T(t)$ and the motor input voltage $v(t)$.

First, we express the relation between the torque produced by the motor m and the motor input voltage v_m. Each motor is controlled in torque and is modeled as follows [Archela *et al.* 2018]:

$$\dot{\tau}_m(t) = -\frac{R_a}{L_a}\tau_m(t) - \frac{K_e K_t}{L_a}\omega_m(t) + \frac{K_t}{L_a}v(t) \tag{7.1}$$

where $\tau_m(t)$ is the motor torque [Nm], $\omega_m(t)$ is the motor angular velocity [rad/s], $v(t)$ is the motor input voltage [V], and R_a is the armature resistance [Ω], L_a is the armature inductance [H], K_e is the back electromotive constant [V/rad/s], K_t is the torque constant [Nm/A].

Then, we express the relation between the motor torque $\tau_m(t)$ and the drive roller torque $\tau_r(t)$. To that end, we consider the subsystem (motor, belt, drive roller) as pictured in Figure 7.2. Assume that (i) there is no loss of transmission via the belt, and (ii) the radius of the drive belt (motor side) is equal to the radius of the driven belt (roller side). From those both assumptions, the drive roller torque $\tau_r(t)$ and velocity $\omega_r(t)$ are equal to the motor torque $\tau_m(t)$ and velocity $\omega_m(t)$, i.e.

$$\tau_r(t) = \tau_m(t), \quad \omega_r(t) = \omega_m(t) \tag{7.2}$$

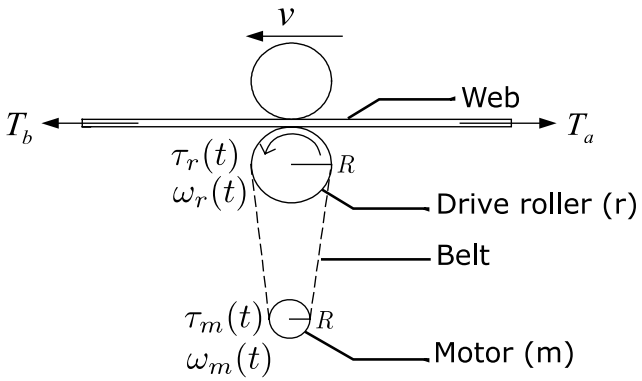


Figure 7.2: Web tension [Chen *et al.* 2004]

Lastly, we express the relation between the drive roller torque $\tau_r(t)$ and the web tension $T(t)$. According to Newton's law of motion, this relation can be expressed as follows:

$$\begin{aligned} J\dot{\omega}_r(t) &= \tau_r(t) + RT_b(t) - RT_a(t) - B\omega_r(t) \\ &= \tau_r(t) + R(T_b(t) - T_a(t)) - B\omega_r(t) \end{aligned} \tag{7.3}$$

where $T_a(t)$, $T_b(t)$ is the web tension of zone a, b respectively [N], and J is the sum of the inertia of motor, belt, drive roller [kg.m^2], B is the viscous friction coefficient [N.m.s].

For an idler roller, equation (7.3) becomes (7.4) since it is not controlled by a motor, and under some other assumptions detailed in [Chen *et al.* 2004].

$$I_t\dot{\omega}_t(t) = R(T_b(t) - T_a(t)) \tag{7.4}$$

where I_t is the rotation inertia [kg.m^2].

7.2.1.2 Web speed $V(t)$

In this subsection, we express the relation between the web tension $T(t)$ and the roller angular velocities. Consider Figure 7.3 where two rollers $r1$, $r2$ transport the web. $S(t)$ represents the web length [m] that depends on the web variation generated by the rollers

$$S(t) = S_0 + \int_{t_0}^{t} R\omega_r 1(\zeta)\,\mathrm{d}\zeta - \int_{t_0}^{t} R\omega_r 2(\zeta)\,\mathrm{d}\zeta \tag{7.5}$$

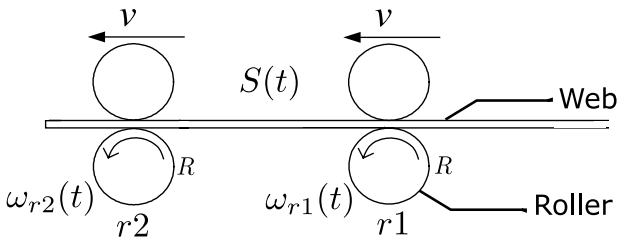


Figure 7.3: Web speed

Then, the web tension at a zone i can be modeled in (7.6).

$$\dot{T}_i(t) = K\dot{S}(t) + D\ddot{S}(t) \tag{7.6}$$

where K and D represent the spring constant [N/m] and the damping constant [Ns/m] of the web, respectively. They are properties of the material.

Hence, (7.6) can be rewritten as

$$\dot{T}(t) = K(R\omega_r 1(t) - R\omega_r 2(t)) + D(R\dot{\omega}_r 1(t) - R\dot{\omega}_r 2(t)) \tag{7.7}$$

Equation (7.7) establishes the relation between the web speed and the web tension.

7.2.1.3 System model

From the relation between the variables in (7.1), (7.3), (7.4), (7.7), the system can be modeled by the set of differential equations in (7.8).

$$\begin{cases} \dot{\tau}_x(t) = -\frac{R_a}{L_a}\tau_x(t) - \frac{K_e K_t}{L_a}\omega_x(t) + \frac{K_t}{L_a}v_x(t) \\ \dot{\tau}_y(t) = -\frac{R_a}{L_a}\tau_y(t) - \frac{K_e K_t}{L_a}\omega_y(t) + \frac{K_t}{L_a}v_y(t) \\ J\dot{\omega}_x(t) = \tau_x(t) - B\omega_x(t) + R(T_1(t) - T_3(t)) \\ J\dot{\omega}_y(t) = \tau_y(t) - B\omega_y(t) + R(T_3(t) - T_2(t)) \\ I_t\dot{\omega}_t(t) = R(T_2(t) - T_1(t)) \\ \dot{T}_1(t) = K(R\omega_t(t) - R\omega_x(t)) + D(R\dot{\omega}_t(t) - R\dot{\omega}_x(t)) \\ \dot{T}_2(t) = K(R\omega_y(t) - R\omega_t(t)) + D(R\dot{\omega}_y(t) - R\dot{\omega}_t(t)) \\ \dot{T}_3(t) = K(R\omega_x(t) - R\omega_y(t)) + D(R\dot{\omega}_x(t) - R\dot{\omega}_y(t)) \end{cases} \tag{7.8}$$

The system in (7.8) can be rewritten in the matrix form in (3.3) with $x(t) = [x_1, x_2, x_3, x_4, x_5, x_6, x_7]^\top = [\tau_x(t), \tau_y(t), \omega_x(t), \omega_y(t), \omega_t(t), \omega_x(t) + \frac{1}{DR}T_1(t), w_y(t) - \frac{1}{DR}T_2(t)]^\top$, where $T_3(t) = -(T_1(t) + T_2(t)$; and $u(t) = [v_x(t), v_y(t)]^\top$. Note that the state $x(t)$ is chosen such that the system is completely controllable [Chen *et al.* 2004].

$$A = \begin{bmatrix} -\frac{R_a}{L_a} & 0 & -\frac{K_e K_t}{L_a} & 0 & 0 & 0 & 0 \\ 0 & -\frac{R_a}{L_a} & 0 & -\frac{K_e K_t}{L_a} & 0 & 0 & 0 \\ \frac{1}{J} & 0 & -\frac{B+2DR^2}{J} & \frac{DR^2}{J} & 0 & \frac{2DR^2}{J} & -\frac{DR^2}{J} \\ 0 & \frac{1}{J} & \frac{DR^2}{J} & -\frac{B+2DR^2}{J} & 0 & -\frac{DR^2}{J} & \frac{2DR^2}{J} \\ 0 & 0 & \frac{DR^2}{I_t} & \frac{DR^2}{I_t} & 0 & -\frac{DR^2}{I_t} & -\frac{DR^2}{I_t} \\ 0 & 0 & \frac{DR^2}{I_t} - \frac{K}{D} & \frac{DR}{I_t} & \frac{K}{D} & -\frac{DR^2}{I_t} & -\frac{DR}{I_t} \\ 0 & 0 & \frac{DR}{I_t} & \frac{DR^2}{I_t} - \frac{K}{D} & \frac{K}{D} & -\frac{DR}{I_t} & -\frac{DR^2}{I_t} \end{bmatrix}, B = \begin{bmatrix} \frac{K_t}{L_a} & 0 \\ 0 & \frac{K_t}{L_a} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, a = 0_7 \tag{7.9}$$

where $R_a = 4.74$, $L_a = 344 \times 10^{-6}$, $K_e = 2.4 \times 10^{-2}$, $K_t = 4.2 \times 10^{-2}$, $R = 2 \times 10^{-2}$, $J = I_t = 8 \times 10^{-4}$, $D = 2$, $K = 3.69 \times 10^2$.

In [Chen *et al.* 2004], the objective is to control the tension and speed of the web defined by $T(t) = DR(-x_3(t) + x_4(t) + x_6(t) - x_7(t))$ and $V(t) = Rx_3(t)$ respectively. To achieve this, the system is equipped with sensors measuring the state $[w_x(t), w_y(t), T_1(t), T_2(t)]$. Note that $x_6(t)$ and $x_7(t)$ can be computed from the measurement of $w_x(t)$, $T_1(t)$, and $w_y(t)$, $T_2(t)$ respectively. Hence, the tension $T(t)$ and the speed $V(t)$ is the service to deliver, and by extension $[w_x(t), w_y(t), T_1(t), T_2(t)]$ are the states that must belong to the service set. In addition, notice that the system in (7.9) is Lyapunov stable.

Consider the input limitation defined by $\mathcal{E}_u(R, \bar{u})$ with

$$R = diag(10^{-2}, 10^{-2}), \quad \bar{u} = 0_2. \tag{7.10}$$

7.2.2 Analysis and prevention of AT1

Let the service set $\mathcal{E}_s(\Xi, \bar{\xi})$ be defined for

$$\begin{aligned} &\Xi = diag(1 \times 10^{-8}, 1 \times 10^{-8}, 1 \times 10^{-2}, 1 \times 10^{-2}, 1 \times 10^{-8}, 2.5 \times 10^{-1}, 2.5 \times 10^{-1}), \\ &\bar{\xi} = 0_7 \end{aligned} \tag{7.11}$$

with $[w_x(t), w_y(t), T_1(t), T_2(t)]$ the state corresponding to the service to deliver, so that it is constrained by $\mathcal{E}_s$, whereas $[\tau_x(t), \tau_y(t), \omega_t(t)]$ is the free state.
Lastly, consider the objective of stealthy aging attacks of AT1 to accelerate the motor aging by controlling them outside of their normal torque levels [Sikanen *et al.* 2018]: $|\tau_x| \geq 1\text{Nm}$, $|\tau_y| \geq 1\text{Nm}$. Hence, the abnormal aging set D_x is defined in (7.12).

$$c_1 = \begin{bmatrix} 1 & 0_{n-1}^\top \end{bmatrix}^\top, \quad c_2 = \begin{bmatrix} -1 & 0_{n-1}^\top \end{bmatrix}^\top, \quad c_3 = \begin{bmatrix} 0 & 1 & 0_{n-2}^\top \end{bmatrix}^\top, \quad c_4 = \begin{bmatrix} 0 & -1 & 0_{n-2}^\top \end{bmatrix}^\top, \tag{7.12}$$

$$\mu = 1 \tag{7.13}$$

From the system model in (7.9), the actuator limitation (7.10), the service set (7.11), we want to verify if the physical integrity of the system, i.e. actuator aging, can be violated by an attacker (RQ1). If it is possible, we want to restrain the control input set $\mathcal{E}_(R, \bar{u})$ into $\hat{\mathcal{E}}_u(\hat{R}, \bar{u})$ to prevent potential stealthy aging attacks of AT1 (RQ2). RQ1 and RQ2 are firstly answered with the invariant set-based method, and then with the convex robust simulation-based method.

7.2.2.1 Invariant set-based method

First-second step: From this first step, we want to quantify the effect induced by stealthy aging attacks of AT1. To do so, Algorithm 2 is applied with the inputs: (7.10), (7.11), (7.9), $\underline{\alpha} = 0$, $\bar{\alpha} = 15$, $\Delta_\alpha = 1 \times 10^{-3}$. It gives in output the smallest constrained invariant state ellipsoid $\mathcal{E}_x$ defined with Q_x and $\bar{x}$ in (7.14). It is obtained for $\alpha = 14.99$, $\beta = 1.1 \times 10^{-3}$, and $\lambda = 3.21 \times 10^{-4}$.

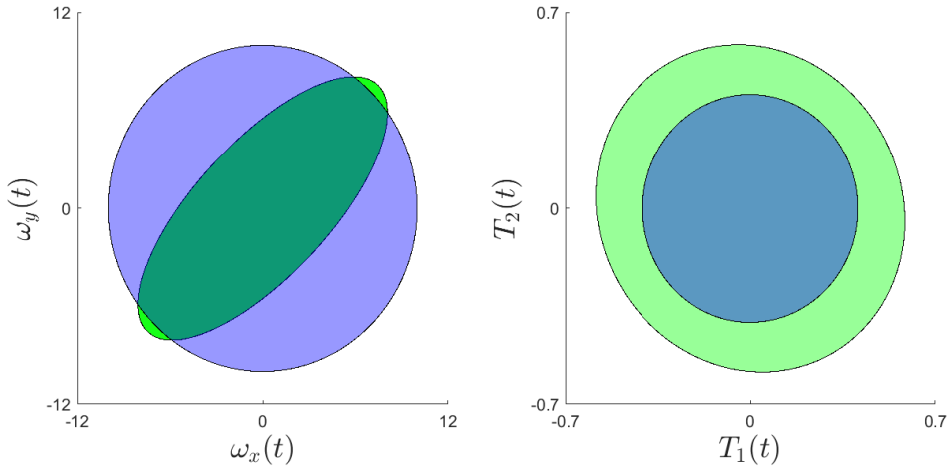
$$Q = 10^{-3} \times \begin{bmatrix} 215.18 & 0.02 & -8.39 & -0.05 & 7.58 & -2.03 & -0.04 \\ 0.02 & 215.18 & -0.05 & -8.39 & 7.58 & -0.04 & -2.03 \\ -8.39 & -0.05 & 93.25 & 0.56 & -83.81 & 22.59 & 0.45 \\ -0.05 & -8.39 & 0.56 & 93.25 & -83.81 & 0.45 & 22.59 \\ 7.58 & 7.58 & -83.91 & -83.81 & 165.49 & -23.37 & -23.37 \\ -2.03 & -0.04 & 22.59 & 0.45 & -23.37 & 13.73 & 0.22 \\ -0.04 & -2.03 & 0.45 & 22.59 & -23.37 & 0.22 & 13.73 \end{bmatrix}, \quad \bar{x} = 0_7 \tag{7.14}$$

The result is shown graphically in Figure 7.4 where the smallest constrained invariant state ellipsoid $\mathcal{E}_x$ is in green, the service ellipsoid $\mathcal{E}_s$ constraining the state is in blue, the abnormal aging set D_x is the red area, and the control input set $\mathcal{E}_u$ is in magenta. As an interpretation, if the control signal $u(t)$ remains inside the control input set $\mathcal{E}_u \forall t \geq t_0$, then the state $x(t)$ remains inside the constrained invariant state ellipsoid $\mathcal{E}_x$ for any initial condition $x(t_0)$ inside this constrained invariant state ellipsoid $\mathcal{E}_x$ and under Assumption 2. From the left- to the right-hand side and from the top to the bottom, the state $[\omega_x(t), \omega_y(t)]$, $[T_1(t), T_2(t)]$, $[V(t), T(t)]$, $[\tau_x(t), \tau_y(t)]$, and the input $[u_x(t), u_y(t)]$ are projected. Notice that the service set has not been drawn for $[\tau_x(t), \tau_y(t)]$ for visibility purpose, but the service set covers the whole space as $\tau_x(t)$ and $\tau_y(t)$ are two free states. One can notice that the state $\tau_x(t)$ and $\tau_y(t)$ can reach the abnormal aging set D_x in red, i.e. $\mathcal{E}_x \cap D \neq \emptyset$. Recall that this observation is indeed computed by Algorithm 3 Thus, stealthy aging attacks can potentially occur if one manipulates the control signal $u(t)$. This answers RQ1, and leads to the next research question RQ2.

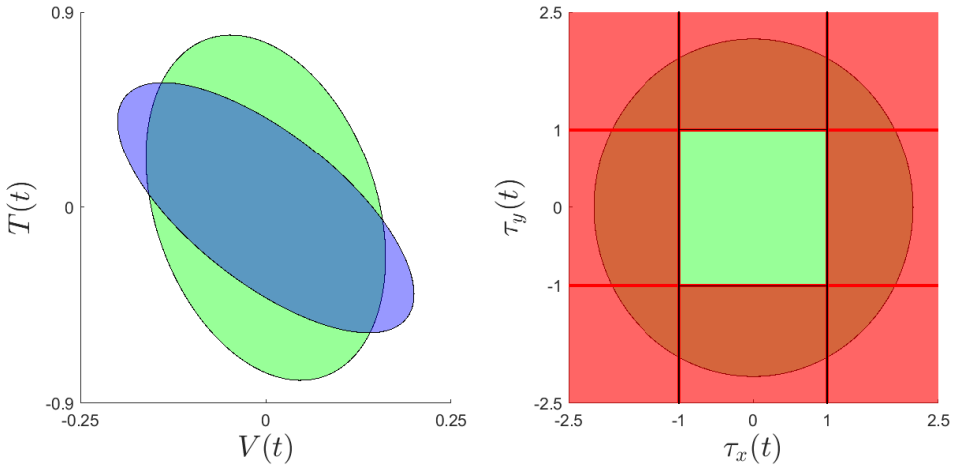
Third step: As stealthy aging attacks are feasible for the given control input set $\mathcal{E}_u$, we want to restrain it such that stealthy aging attacks cannot occur. Algorithm 4 is applied with the inputs: $R_0 = I_m$, $\underline{\delta} = 1 \times 10^{-2}$, $tol = 1 \times 10^{-3}$. It converges, with respect to the tolerance tol, to the largest control input set $\hat{\mathcal{E}}_u(\hat{R}, \bar{u})$ with $\hat{R} = \delta R_0$ defined for $\delta = 4.9 \times 10^{-2}$

$$\hat{R} = 10^{-2} \times \begin{bmatrix} 4.09 & 0 \\ 0 & 4.09 \end{bmatrix}, \quad \bar{u} = 0_2 \tag{7.15}$$

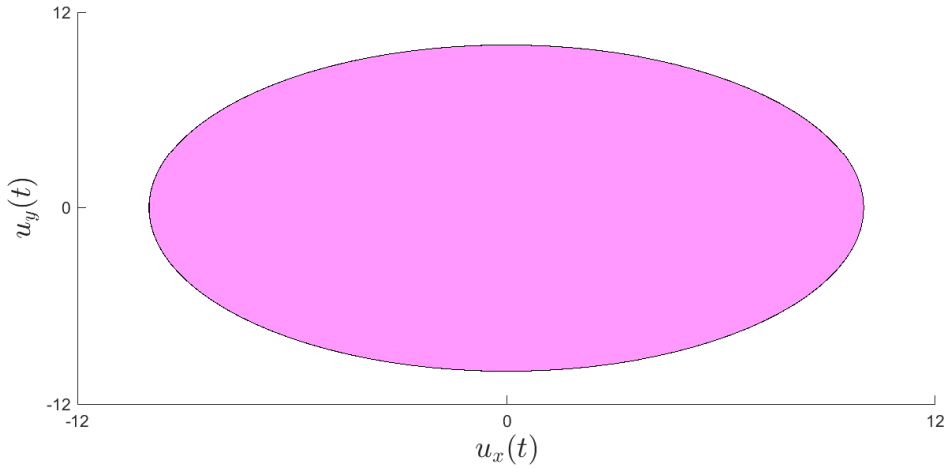
such that the smallest constrained invariant state ellipsoid $\mathcal{E}_x$ obtained for $\alpha = 14.99$ over the interval $\bar{\alpha} = 15$ with a step $\Delta_\alpha = 0.01$ with $\beta = 13.37$, $\lambda = 1.62$



(a) Figure A



(b) Figure B



(c) Figure B

Figure 7.4: AT1 for the film mfg. system with invariant-set based method - RQ1 - constrained invariant state set $\mathcal{E}_x$ (green), service set $\mathcal{E}_s$ (blue) covering the whole space for $[\tau_x,\tau_y]$ (not drawn), abnormal aging set D_x (red), control input set $\mathcal{E}_u$ (magenta)

$$Q = 10^{-3}\times\begin{bmatrix} 1013.87 & 0.02 & -19.59 & -0.23 & 16.17 & -2.16 & -0.09 \\ 0.02 & 1013.87 & -0.23 & -19.59 & 16.17 & -0.09 & -2.16 \\ -19.59 & -0.23 & 218.63 & 2.57 & -178.49 & 23.94 & 0.99 \\ -0.23 & -19.59 & 2.57 & 218.63 & -178.49 & 0.99 & 23.94 \\ 16.17 & 16.17 & -178.49 & -178.49 & 338.80 & -26.18 & -26.18 \\ -2.16 & -0.09 & 23.94 & 0.99 & -26.18 & 9.95 & 0.07 \\ -0.09 & -2.16 & 0.99 & 23.94 & -26.18 & 0.07 & 9.95 \end{bmatrix}, \quad \bar{u} = 0_7 \tag{7.16}$$

does not intersect with the abnormal aging set D_x.

The result is graphically shown in Figure 7.5 where the smallest constrained invariant state ellipsoid $\mathcal{E}_x$ is in green, the service ellipsoid $\mathcal{E}_s$ constraining the state is in blue, the abnormal aging set D_x is the red area, and the restrained control input set $\hat{\mathcal{E}}_u$ is in magenta. As we can observe in Figure 7.5, the abnormal aging set is guaranteed unreachable for $u(t) \in \hat{\mathcal{E}}_u$, $\forall t \geq t_0$. Therefore, stealthy aging attacks are prevented by restraining the control input $u(t)$ to the restrained input set.

7.2.2.2 Convex robust simulation-based method

Notice that in this application, we have considered $\tilde{Q}(t)$ MSOS of degree 4 ($d = 2$); $Z(t)$, $X(t)$ are matrices (MSOS of degree 0); $\alpha(t)$ SOS of degree 4 ($d = 2$); and $\beta(t)$ an unknown scalar (SOS of degree 0).

First-second step:

The analysis of stealthy aging attacks of AT1 is now studied on the finite time horizon with the convex robust simulation-based method. Instead of stopping the simulation once a constrained bounding ellipsoid at time $t = j\Delta_t$ in Algorithm 6 hits with the abnormal aging set, we stop it once the bounding ellipsoids converge, in this case for $t_f = 1s$. We apply Algorithm 6 with the inputs: $Q(t_0) = diag(10, 10, 10, 10, 10, 10, 10)$, $\bar{x}(t_0) = [0, 0, 0, 0, 0, 0, 0]^\top$, $\Delta_t = 0.1ms$, $t_f = 1s$. The result is graphically shown in Figure 7.6, where the final bounding state ellipsoid $\mathcal{E}_x(t_f)$ is in green, the bounding state ellipsoids $\mathcal{E}_x(t_i)$ with $t_i = \{t_0, ..., t_f - \Delta_t\}$ are not filled, i.e. only their bound is drawn, the service ellipsoid $\mathcal{E}_s$ constraining the state is in blue, the abnormal aging set D_x is the red area, and the input ellipsoid $\mathcal{E}_u$ is in magenta. Recall that the state starting from a bounding state ellipsoid $\mathcal{E}_x(t_i)$ remains inside $\mathcal{E}_x(t_i + \Delta_t)$ for $u(t) \in \mathcal{E}_u$ and for $x(t) \in \mathcal{E}_s$, $\forall t \in [t_i, t_i + \Delta_t]$ under Assumption 1.

The bounding state ellipsoid hits for the first time the abnormal aging set at $t_\cap = 9.5ms$, i.e. $\mathcal{E}_x(9.5\text{ms}) \cap D_x \neq \emptyset$ that corresponds to the $j = 95^{th}$ iterations in Algorithm 6. From this analysis, we can conclude that the physical integrity of the system can be violated over the time interval $[0, 1s]$ by an attacker manipulating the control signal $u(t)$. Thus, stealthy aging attacks are feasible which answers RQ1, and leads to the prevention step (third step).

Third step: In the third step, we want to restrain the control input set $\mathcal{E}_u$ such that stealthy aging attacks are not feasible. To do so, we apply Algorithm 7 with the inputs: $R_0 = I_2$, $\underline{\delta} = 1 \times 10^{-2}$, $tol = 1 \times 10^{-3}$, $Q(t_0) = diag(10, 10, 10, 10, 10, 10, 10)$, $\bar{x}(t_0) =$

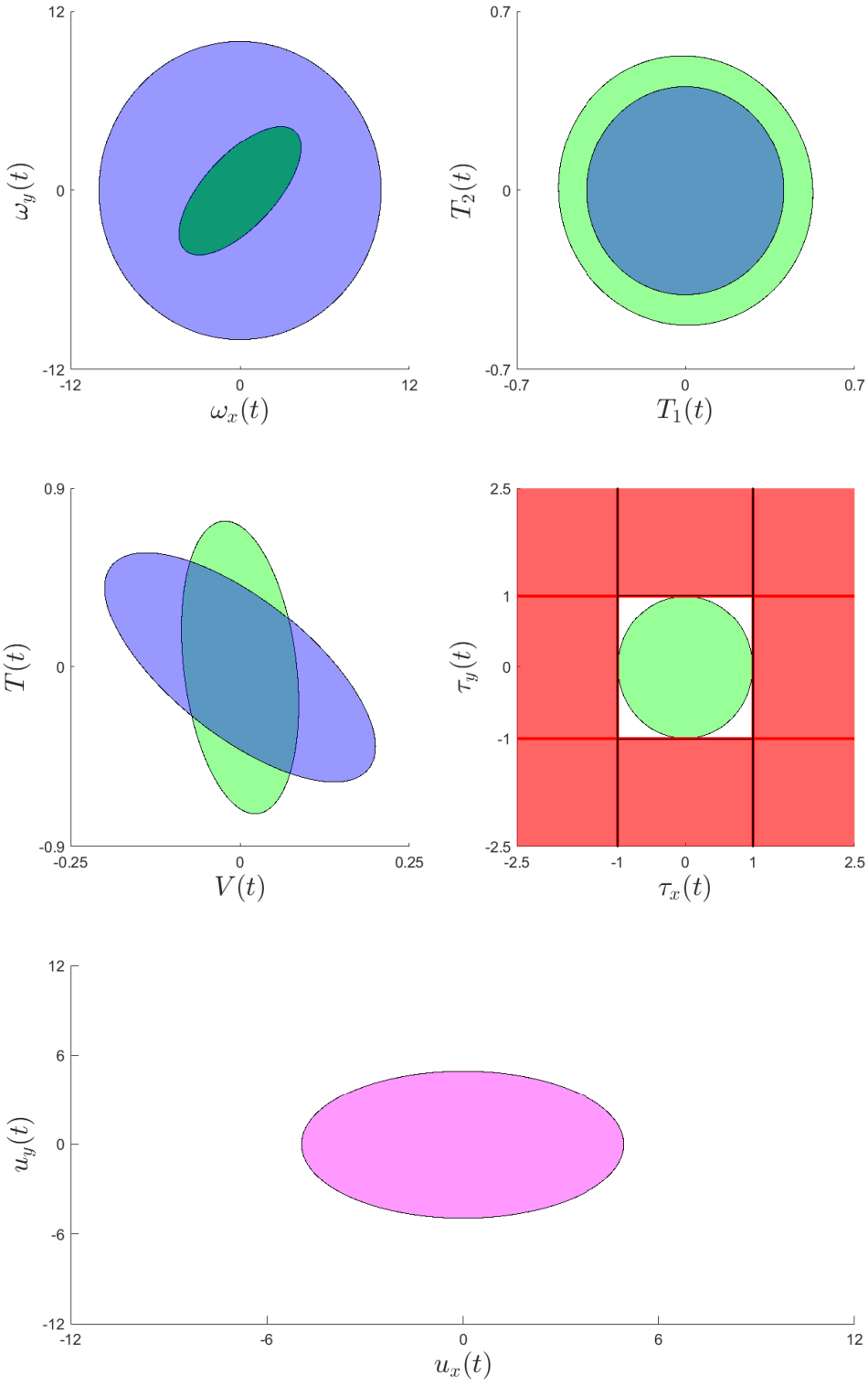


Figure 7.5: AT1 for the film mfg. system with invariant-set based method - RQ2 - constrained invariant state set $\mathcal{E}_x$ (green), service set $\mathcal{E}_s$ (blue) covering the whole space for $[\tau_x,\tau_y]$ (not drawn), abnormal aging set D_x (red), restrained control input set $\hat{\mathcal{E}}_u$ (magenta)

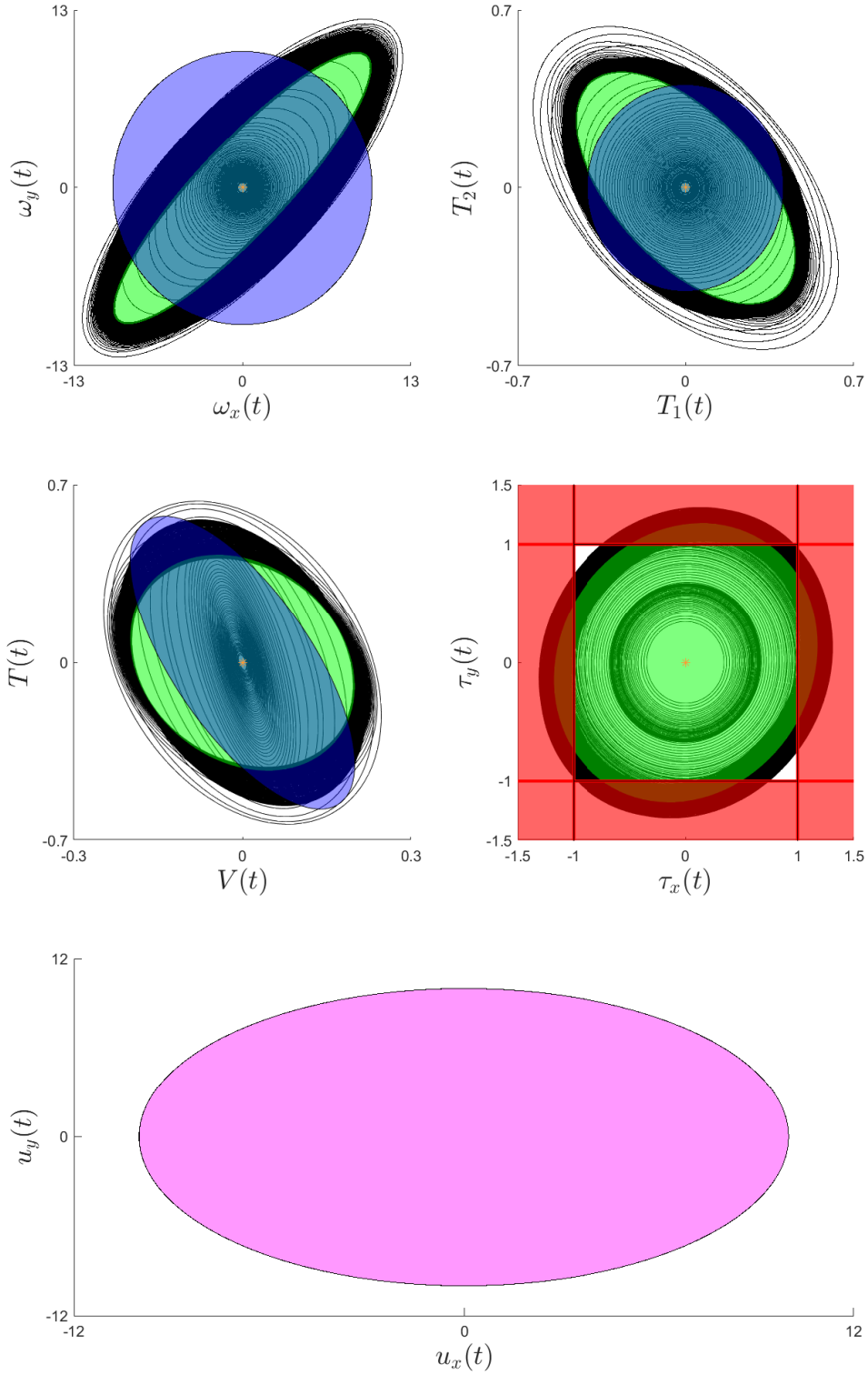


Figure 7.6: AT1 for the film mfg. system with robust convex simulation-based method - RQ1 - Bounding state ellipsoids $\mathcal{E}_x(t_i)$ at each 0.1ms from initial set $\mathcal{E}_x(t_0)$ (orange star) and final set $\mathcal{E}_x(t_f)$ (green), abnormal aging set D_x (red), control input set $\mathcal{E}_u$ (magenta)

$[0,0,0,0,0,0,0]^\top$, $\Delta_t = 0.1ms$, $t_f = 1s$. The largest restrained control input set $\hat{\mathcal{E}}_u(\hat{R}, \bar{u})$ is given from the output $\delta = 0.0168$ of Algorithm 7 and computed as $\hat{R} = \delta R_0$

$$\hat{R} = 10^{-2} \times \begin{bmatrix} 1.68 & 0 \\ 0 & 1.68 \end{bmatrix}, \quad \bar{u} = 0_2 \tag{7.17}$$

The result is graphically shown in Figure 7.7 where the bounding ellipsoids $\mathcal{E}_x(t_i)$ are the ones for the restrained input set $\hat{\mathcal{E}}_u$. We can observe that none of the bounding ellipsoids intersect with the abnormal aging set, so stealthy aging attacks are prevented for any input $u(t) \in \hat{\mathcal{E}}_u$.

7.2.3 Analysis of AT2

In this subsection, stealthy aging attacks of AT2 is analyzed. The system model in (7.9) is discretized with the Tustin method with a sampling time $Te = 1ms$.

Consider the attacker aims to maximize the thermomechanical stresses [Sikanen *et al.* 2018] (e.g. heat) due to the torque square in the motor x. Hence, the degradation function is expressed as $g_k = \tau_{x,k}^2$. Recall that the attacker can manipulate the control signal, and it knows the desired service at each time instant k. Thus and before analyzing the attack, a preliminary computation is required to get the normal signals (norm) that a controller should produce in the attack-free case. The signal of interest in the normal signals are the ones that belong to the service set, so that X_s can be defined. Then, the attack input signal u_k is computed with the SLP method to maximize the cost degradation function while keeping the service states $[x_3, x_4, x_6, x_7]$ inside X_s given for a $\delta = 10^{-4}$ error around the normal service. At the 28^{th} iteration of the SLP, the cost degradation function converges as shown in Fig. 7.8.

7.3 Satellite reorientation maneuver control

In this section, we propose to show the relevance of the proposed methods for AT1 and AT2 upon a satellite reorientation maneuver control.

7.3.1 System description

Consider an attitude controlled satellite, featuring four reaction wheels as actuators [Wie 2008]. The control objective is to slightly re-orientate it at the orientation reference values. The Euler angles α, β, γ [rad] characterize the orientation of the coordinate system. The satellite is equipped with four reaction wheels w_i with $i = \{1, 2, 3, 4\}$ having the same characteristics: an inertia J_w. Each of them is controlled by an electrical motor m_i, having the same characteristics: an armature resistance R_a, an electric constant k_e, a torque constant k_t. As in the previous application, we will not detail the modeling given in [Wie 2008], but explain the physical relations between the variables.

7.3.1.1 Euler angles

The total angular momentum of the satellite is defined in (7.18).

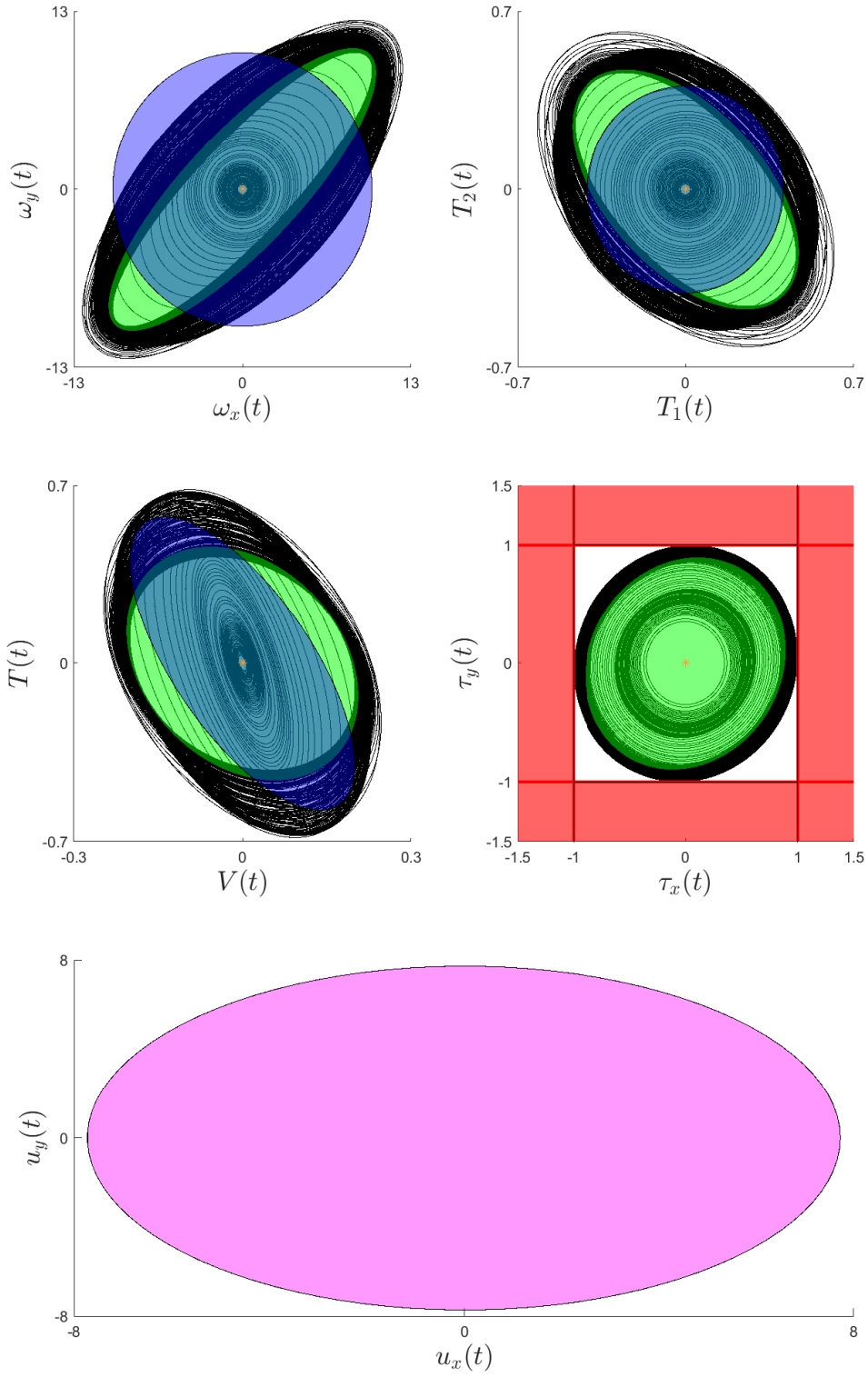


Figure 7.7: AT1 for the film mfg. system with robust convex simulation-based method - RQ2 - Bounding state ellipsoids $\mathcal{E}_x(t_i)$ at each 0.1ms from initial set $\mathcal{E}_x(t_0)$ (orange star) and final set $\mathcal{E}_x(t_f)$ (green), abnormal aging set D_x (red), restrained control input set $\hat{\mathcal{E}}_u$ (magenta)

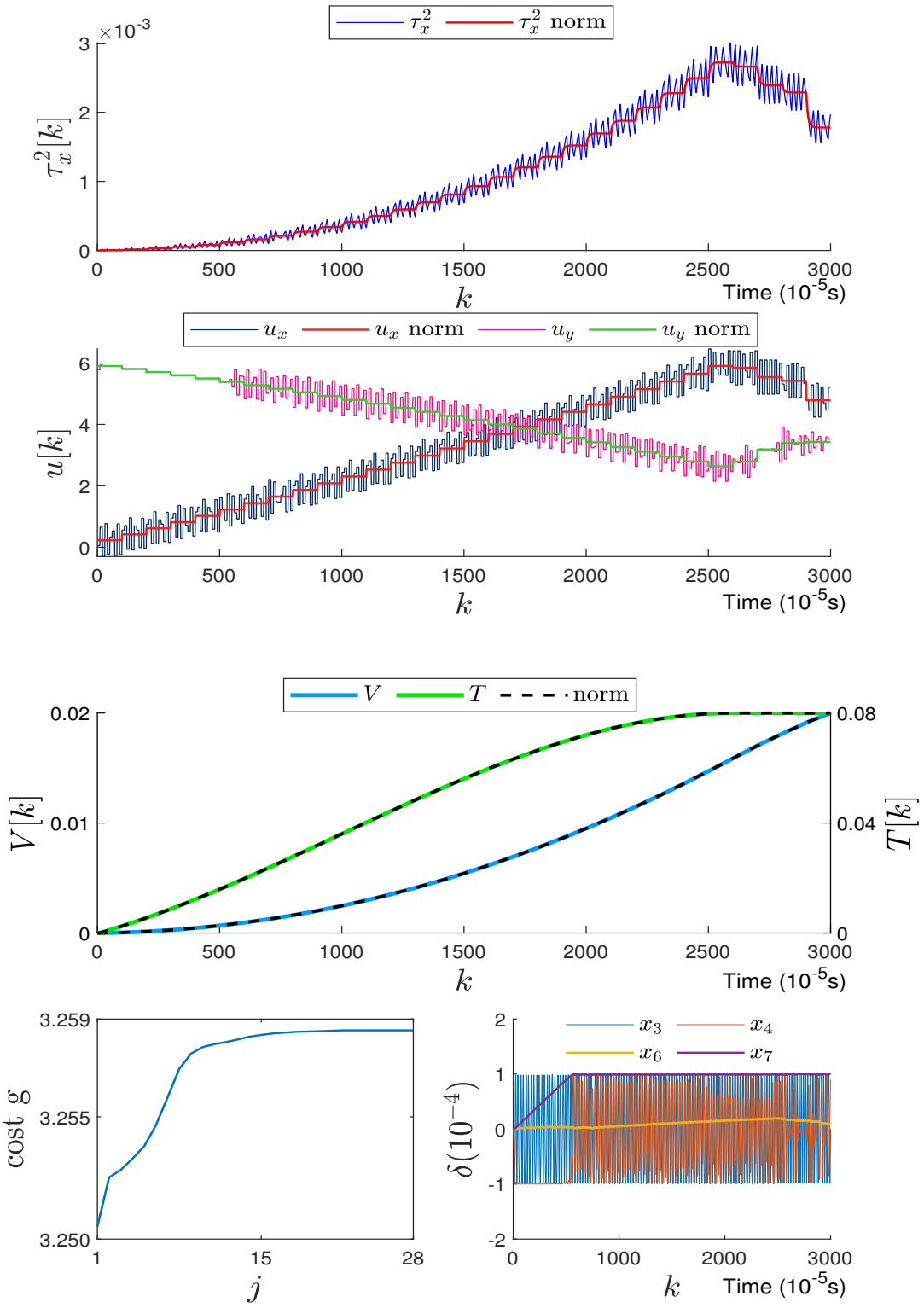


Figure 7.8: AT2 for the film mfg. control: at the 28$^{\text{th}}$ iteration, the cost degradation function is maximized (1^{st} from bottom) with the degradation function (1^{st} from top) for the optimal attack input signal (2^{nd} from top) while delivering the desired service (2^{nd} from bottom)

$$H_s(t) = I_s\omega_s(t) + h_w(t) \tag{7.18}$$

where $\omega_s(t) = [\dot{\alpha}(t), \dot{\beta}(t), \dot{\gamma}(t)]^\top$ is the satellite angular velocity in body frame with $\dot{\alpha}$, $\dot{\beta}$, $\dot{\gamma}$ the Euler angular rates [rad/s], $h_w(t)$ is the angular momentum of the reaction wheels, and I_s is the inertia matrix of the satellite.
The angular momentum of the reaction wheels $h_w(t)$ can be described in body frame in (7.19).

$$h_w(t) = L_w I_w \omega_w(t) \tag{7.19}$$

where $\omega_w(t) = [\omega_{w1}(t), \omega_{w2}(t), \omega_{w3}(t), \omega_{w4}(t)$ is the angular velocity of the reaction wheels, and L_w is the reaction wheel distribution matrix, $I_w = \text{diag}(I_{w1}, I_{w2}, I_{w3}, I_{w4})$ is the reaction wheel diagonal inertia matrix.

Assuming the reaction wheels have a pyramid configuration, then the moments actuated on the satellite are defined by

$$T_s(t) = L_w T_w(t), \text{with } L_w = \begin{bmatrix} -\sqrt{3}/3 & \sqrt{3}/3 & \sqrt{3}/3 & -\sqrt{3}/3 \\ -\sqrt{3}/3 & -\sqrt{3}/3 & \sqrt{3}/3 & \sqrt{3}/3 \\ \sqrt{3}/3 & \sqrt{3}/3 & \sqrt{3}/3 & \sqrt{3}/3 \end{bmatrix} \tag{7.20}$$

where $T_s(t) = [\tau_\alpha(t), \tau_\beta(t), \tau_\gamma(t)]^\top$ the torque actuating on each axes of the body frame, and $T_w(t) = [\tau_{w1}(t), \tau_{w2}(t), \tau_{w3}(t), \tau_{w4}(t)]^\top$ is the torque generated by each reaction wheel wi. From the moment equations of Euler represented by the angular momentum rate with respect to the body frame, the dynamics of the satellite can be described in (7.21), assuming no disturbance torque.

$$\dot{H}_s(t) = -\omega_s(t) \times H_s(t) \tag{7.21}$$

Then, we neglect the term $\omega_s(t) \times H_s(t)$, which yields $\dot{H}_s(t) = 0$.
The first time derivative of (7.18) yields to

$$\dot{H}_s(t) = I_s\dot{\omega}_s(t) + \dot{h}_w(t) \tag{7.22}$$

Hence, (7.21) can be rewritten from the first time (7.22) as follows:

$$I_s\dot{\omega}_s(t) = -\dot{h}_w(t) \tag{7.23}$$

Each reaction wheel produces an angular momentum which is transferred to the satellite, so $\dot{h}_w(t) = T_s(t) = L_w T_w(t)$.

7.3.2 Reaction wheel torque and motor angular velocity

In this subsection, we express the relation between the reaction wheel torque $\tau_{wn}(t)$ and the motor input voltage $v_{mn}(t)$. This is expressed for a motor m_n and reaction wheel w_n as is validate for all the motors and reaction wheels of the application.

As the reaction wheel is fixed to the rotor shaft, then the reaction wheel and the motor have same angular velocity and acceleration, i.e. $\omega_{wn}(t) = \omega_{mn}(t)$, $\dot{\omega}_{wn}(t) =$

$omega_{mn}(t)$. Consider a motor controlled in torque similar to the previous application. From [Archela *et al.* 2018], the armature current can be expressed in (7.24).

$$L_a\dot{i}_{mn}(t) = K_e\omega_{mn}(t) - R_a i_{mn}(t) + v_{mn}(t) \tag{7.24}$$
$$\tau_{mn}(t) = K_t i_{mn}(t) \tag{7.25}$$

As the dynamic of the current is very fast compared to the satellite dynamic, the current dynamic is neglected. Hence equation (7.24) is restated at the equilibrium state in (7.26).

$$i_{mn}(t) = \frac{K_e}{R_a}\omega_{mn}(t) + \frac{1}{R_a}v_{mn}(t) \tag{7.26}$$

The moment balance of the motor shaft is given in (7.27).

$$I_{mn}\dot{\omega}_{mn}(t) = \tau_{mn}(t) - B\omega_{mn}(t) - \tau_{wn}(t) \tag{7.27}$$

where $\tau_{wn}(t) = I_{wn}\dot{\omega}_{wn}(t)$ as the load torque is the reaction wheel fixed on the rotor shaft, and B is the viscous friction coefficient [N.m.s] of the motor, I_{mn} is the moment of inertia of the motor.

Thus, as $\dot{\omega}_{wn}(t) = \dot{\omega}_w(t)$, (7.27) can be rewritten

$$(I_{mn} + I_{wn})\dot{\omega}_{wn}(t) = \tau_{mn}(t) - B\omega_{mn}(t) \tag{7.28}$$

It is reasonable to assume that $I_{wn} >> I_{mn}$, thus equation (7.28) becomes (7.29).

$$I_{wn}\dot{\omega}_{wn}(t) \approx \tau_{mn}(t) - B\omega_{mn}(t) \tag{7.29}$$

Lastly, by substituting equation (7.26) into (7.29) with (7.25), the motor angular velocity $\omega_{wn}(t)$ can be expressed in (7.30).

$$I_{wn}\dot{\omega}_{mn}(t) \approx (\frac{K_t K_e}{R_a} - B)\omega_{mn}(t) + \frac{K_t}{R_a}v_{mn}(t) \tag{7.30}$$

Then, we express the satellite angular acceleration $\dot{\omega}_s(t)$ in by substituting (7.29) into (7.23).

$$I_s\dot{\omega}_s(t) = -L_w I_4(\frac{K_t K_e}{R_a} - B)\omega_m(t) - L_w I_4\frac{K_t}{R_a}v_m(t) \tag{7.31}$$

7.3.3 System model

From the relation between the variables in (7.30), (7.31), the system can be modeled by the set of differential equations in (7.32).

$$\begin{cases} \dot{\omega}_m(t) = I_{wn}^{-1}(\frac{K_t K_e}{R_a} - B)\omega_{mn}(t) + I_{wn}^{-1}\frac{K_t}{R_a}v_{mn}(t) \\ \dot{\omega}_s(t) = -I_s^{-1}L_w I_4(\frac{K_t K_e}{R_a} - B)\omega_m(t) - I_s^{-1}L_w I_4\frac{K_t}{R_a}v_m(t) \end{cases} \tag{7.32}$$

The system in (7.32) can be rewritten in the matrix form in (3.3) with $x(t) = [x_1, x_2, x_3, x_4, x_5, x_6, x_7]^\top = [\omega_1(t), \omega_2(t), \omega_3(t), \omega_4(t), \alpha(t), \beta(t), \gamma(t), \dot{\alpha}(t), \dot{\beta}(t), \dot{\gamma}(t)]\top$; and $u(t) = [v_1(t), v_2(t), v_3(t), v_4(t)]\top$.

$$A = \begin{bmatrix} I_{wn}^{-1}(\frac{K_t K_e}{R_a} - B) & 0_{4,3} & 0_{4,3} \\ 0_{3,4} & 0_{3,3} & 1_{3,3} \\ -I_s^{-1} L_w I_4 (\frac{K_t K_e}{R_a} - B) & 0_{3,3} & 0_{3,3} \end{bmatrix}, B = \begin{bmatrix} I_{wn}^{-1} \frac{K_t}{R_a} \\ 0_{3,4} \\ -I_s^{-1} L_w I_4 \frac{K_t}{R_a} \end{bmatrix}, a = 0 \tag{7.33}$$

where $J_{wn} = 2.03 \times 10^{-6}$, $R_a = 3$, $K_e = 4.2 \times 10^{-3}$, $K_t = 4.2 \times 10^{-3}$, $B = 2.65 \times 10^{-6}$, $J_s = \text{diag}(4.94 \times 10^{-2}, 4.93 \times 10^{-2}, 2.01 \times 10^{-2})$.

The control objective is to control the Euler angles $[\alpha(t), \beta(t), \gamma(t)]$. To achieve this, the system is equipped with sensors measuring the Euler angular rates $[\dot{\alpha}(t), \dot{\beta}(t), \dot{\gamma}(t)]$. Hence, $[\alpha(t), \beta(t), \gamma(t)]$ is the service to deliver, and $[\dot{\alpha}(t), \dot{\beta}(t), \dot{\gamma}(t)]$ is how the service is perceived to be delivered by the control system. Thus, $[\alpha(t), \beta(t), \gamma(t), \dot{\alpha}(t), \dot{\beta}(t), \dot{\gamma}(t)]$ are the states that are constrained by the service set.

Lastly, consider the control input limitation defined by $\mathcal{E}_u(R, \bar{u}$ with

$$R = diag(4 \times 10^{-2}, 4 \times 10^{-2}, 4 \times 10^{-2}, 4 \times 10^{-2}), \quad \bar{u} = 0_4. \tag{7.34}$$

7.3.4 Analysis and prevention of AT1

Let the service set $\mathcal{E}_s(\Xi, \bar{\xi})$ be defined for

$$\Xi = diag(1 \times 10^{-6}, 1 \times 10^{-6}, 1 \times 10^{-6}, 1 \times 10^{-6}, 4, 4, 4, 4, 4, 4), \quad \bar{\xi} = 0_1 0 \tag{7.35}$$

with $[\alpha(t), \beta(t), \gamma(t), \dot{\alpha}(t), \dot{\beta}(t), \dot{\gamma}(t)]$ is the state corresponding to the service to deliver, so that it is constrained by $\mathcal{E}_s$ whereas $[\omega_1(t), \omega_2(t), \omega_3(t), \omega_4(t)]$ is the free state. Actually, the free state is constrained due to their physical limitation: the maximum angular velocity. Notice that the system is not Lyapunov stable, but the unstable states are constrained by $\mathcal{E}_s$.

Lastly, consider the objective of stealthy aging attacks of AT1 to accelerate the satellite aging by controlling the motors outside of their normal current levels to over-consume power. The abnormal current levels are defined as $|i_1| \geq 2$A, $|i_2| \geq 2$A, $|i_3| \geq 2$A, $|i_4| \geq 2$A. As the current dynamic has been neglected, we express the abnormal aging set D_x from (7.26). In (7.26), the current at the steady-state is expressed from the motor angular velocity and the input voltage. Hence, instead of expressing D_x in terms of $x(t)$, we express it with an extended vector $x_{ext}(t) = [x(t)^\top u(t)^\top]^\top$ such that c becomes $c_{ext} = [c_x \, c_u]$ where c_x, c_u express the linear combination with $x(t)$ and $u(t)$, respectively. Hence, the abnormal aging set D_x is expressed from (7.26) as follows

$$c_{ext,1} = \begin{bmatrix} -a_1 & 0_{n-1} & a_2 & 0_{m-1}^\top \end{bmatrix}^\top, \quad c_{ext,2} = -c_{ext,1} \tag{7.36}$$

$$c_{ext,3} = \begin{bmatrix} 0 & -a_1 & 0_{n-2} & 0 & a_2 & 0_{m-2}^\top \end{bmatrix}^\top, \quad c_{ext,4} = -c_{ext,3} \tag{7.37}$$

$$c_{ext,5} = \begin{bmatrix} 0 & 0 & -a_1 & 0_{n-3} & 0 & 0 & a_2 & 0_{m-3}^\top \end{bmatrix}^\top, \quad c_{ext,6} = -c_{ext,5} \tag{7.38}$$

$$c_{ext,7} = \begin{bmatrix} 0 & 0 & 0 & -a_1 & 0_{n-4} & 0 & 0 & 0 & a_2^\top \end{bmatrix}^\top, \quad c_{ext,8} = -c_{ext,7} \tag{7.39}$$

$$\mu = 2 \tag{7.40}$$

Because the system is not completely controllable, it is transformed with the transformation matrix Z such that $x(t) = Z\breve{x}(t)$, with $\breve{x}(t) = [\breve{x}_1, \breve{x}_2, \breve{x}_3, \breve{x}_4, \breve{x}_5, \breve{x}_6, \breve{x}_7]^\top$ the new (controllable) state, and $\breve{\mathcal{E}}_s$, $\breve{D}$ are the transformed sets of $\mathcal{E}_s$, D_x respectively.

In the following, we apply the robust convex simulation-based method on the system to analyze and prevent stealthy aging attacks of AT1. Since the system is not Lyapunov stable, the invariant set-based method cannot be applied.

7.3.4.1 Robust convex simulation-based method

Notice that in this application, we have considered $\tilde{Q}(t)$ MSOS of degree 6 ($d = 3$); $Z(t)$, $X(t)$ are matrices (MSOS of degree 0); $\alpha(t)$ SOS of degree 6 ($d = 3$); and $\beta(t)$ an unknown scalar (SOS of degree 0).

The analysis of stealthy aging attacks of AT1 on the finite time horizon is studied. Similarly as previously, we stop the simulation once the bounding ellipsoids converge. In this case, it converges at $t_f = 4s$. Algorithm 6 is applied with the inputs: $\breve{Q}(t_0) = diag(100, 100, 100, 100, 100, 100, 100)$, $\breve{\bar{x}}(t_0) = [0, 0, 0, 0, 0, 0, 0]^\top$, $\Delta_t = 0.1s$, $t_f = 4s$. The result is graphically shown in Figure 7.9, where the final bounding state ellipsoid $\mathcal{E}_x(t_f)$ is in green, the bounding state ellipsoids $\mathcal{E}_x(t_i)$ with $t_i = \{t_0, ..., t_f - \Delta_t\}$ are not filled, i.e. only their bound is drawn, the service ellipsoid $\mathcal{E}_s$ constraining the state is in blue, the abnormal aging set D_x is the red area, and the input ellipsoid $\mathcal{E}_u$ is in magenta.

The bounding ellipsoid hits for the first time the abnormal aging set at $t_\cap = 100ms$, i.e. $\mathcal{E}_x(100\text{ms}) \cap D_x \neq \emptyset$ that corresponds to the $j = 1^{st}$ iteration in Algorithm 6. Therefore, the physical integrity of the system can be violated over the time interval $[0, 4s]$ by an attacker manipulating the control signal $u(t)$. Thus, stealthy aging attacks are feasible which answers RQ1, and leads to the prevention step (third step).

Third step: In the third step, we want to restrain the control input set $\mathcal{E}_u$ such that stealthy aging attacks are not feasible. Algorithm 7 is applied with the inputs: $R_0 = I_2$, $\underline{\delta} = 1 \times 10^{-2}$, $tol = 1 \times 10^{-2}$, $\breve{Q}(t_0) = diag(100, 100, 100, 100, 100, 100, 100)$, $\breve{\bar{x}}(t_0) = [0, 0, 0, 0, 0, 0, 0]^\top$, $\Delta_t = 0.1ms$, $t_f = 4s$. The largest restrained control input set $\hat{\mathcal{E}}_u(\hat{R}, \bar{u})$ is given from the output $\delta = 0.1338$ of Algorithm 7 and computed as $\hat{R} = \delta R_0$

$$\hat{R} = 10^{-1} \times \begin{bmatrix} 1.34 & 0 & 0 & 0 \\ 0 & 1.34 & 0 & 0 \\ 0 & 0 & 1.34 & 0 \\ 0 & 0 & 0 & 1.34 \end{bmatrix}, \quad \bar{u} = 0_4 \tag{7.41}$$

The result is graphically shown in Figure 7.10 where the bounding ellipsoids $\mathcal{E}_{\breve{x}}(t_i)$ are the ones for the restrained input set $\hat{\mathcal{E}}_u$. We can observe that none of the bounding ellipsoids intersect with the abnormal aging set, so stealthy aging attacks are prevented for any input $u(t) \in \hat{\mathcal{E}}_u$.

Remark 14. *In Figure 7.9 and 7.10, we can observe that the bounding state ellipsoids grows until they are constrained by the service set. This comes from the stability of the dynamical system: as it is not Lyapunov stable, then the state $\breve{x}(t)$ does not return to the equilibrium state after a perturbation, i.e. the control input $u(t)$ in this case. A physical interpretation is given as follows. By controlling the motors, the satellite moves*

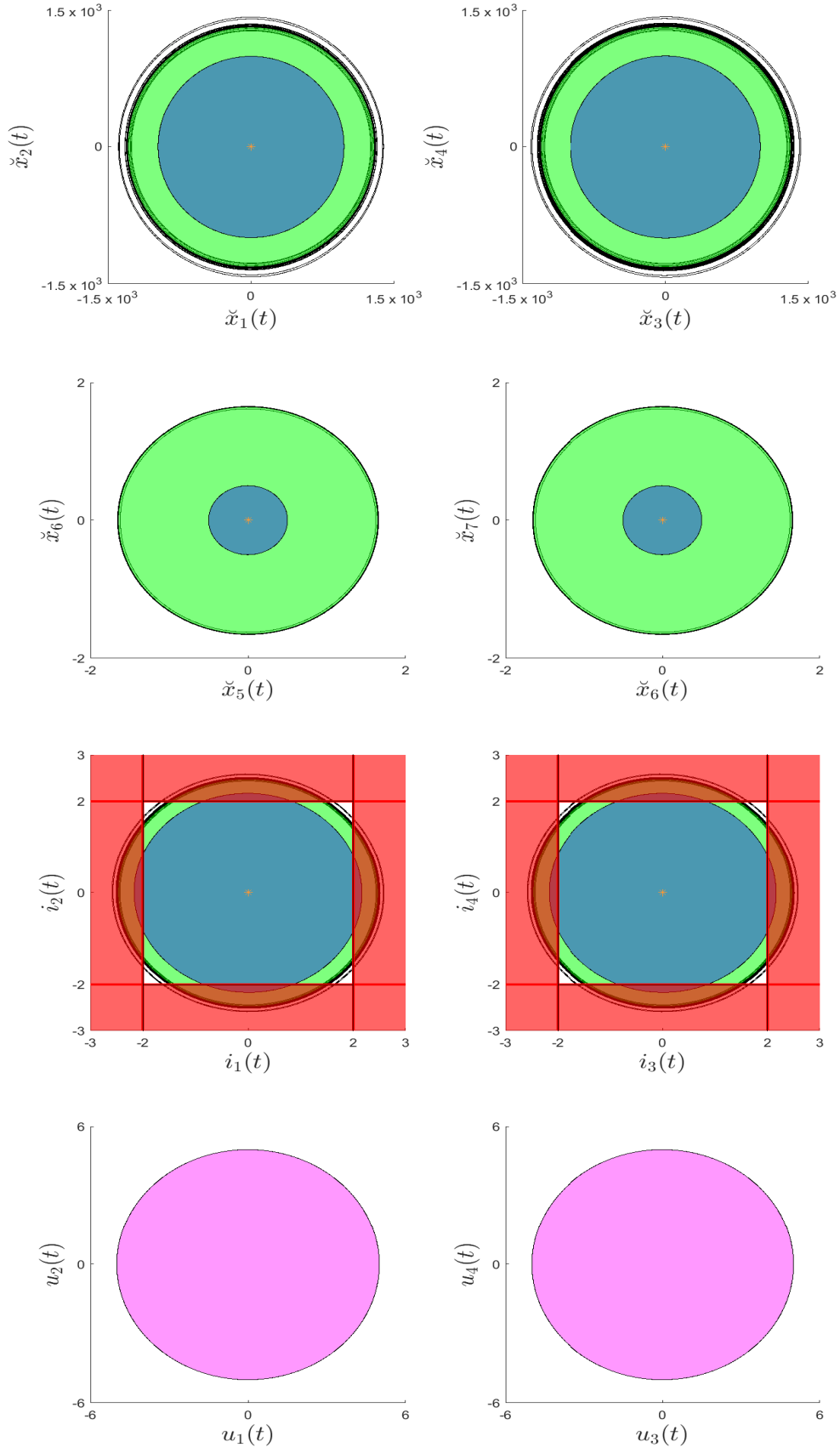


Figure 7.9: AT1 for the satellite with robust convex simulation-based method - RQ1 - Bounding state ellipsoids $\mathcal{E}_x(t_i)$ at each 0.1s from initial set $\mathcal{E}_x(t_0)$ (orange star) and final set $\mathcal{E}_x(t_f)$ (green), abnormal aging set D_x (red), control input set $\mathcal{E}_u$ (magenta)

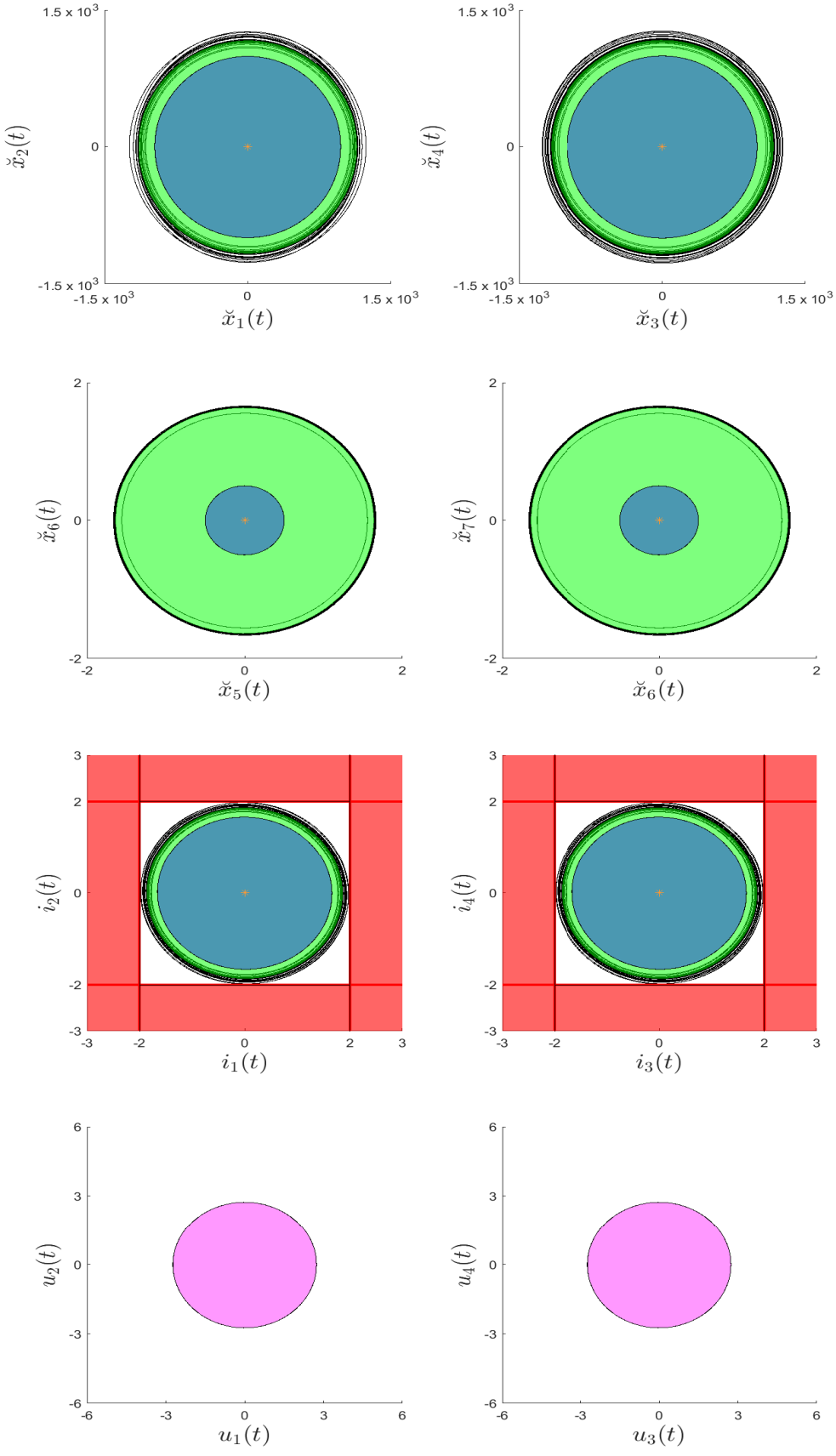


Figure 7.10: AT1 for the satellite with robust convex simulation-based method - RQ2 - Bounding state ellipsoids $\mathcal{E}_x(t_i)$ at each 0.1s from initial set $\mathcal{E}_x(t_0)$ (orange star) and final set $\mathcal{E}_x(t_f)$ (green), abnormal aging set D_x (red), restrained control input set $\hat{\mathcal{E}}_u$ (magenta)

according to the Euler angles. Recalling no external torque actuating on the satellite has been considered and another term in the dynamics of the satellite, then $\dot{H}_s(t) = 0$. *In other words, we do not considered the dynamics of the satellite. Thus, the Euler angular rate never returns to* 0, *unless if a motor actuates oppositely. That is why the bounding state ellipsoid grows as the state always grow, but they are constrained by the service set.*

7.3.5 Analysis of AT2

In this subsection, we propose to analyze stealthy aging attacks of AT2. The system model in (7.33) is discretized with the Tustin method with a sampling time $Te = 10ms$.

Consider the attacker aims to maximize the power consumption of the satellite. The degradation function is expressed as $g_k = \sum_{i=1}^{4} p_{i,k}$, $p_{i,k} = i_{i,k}u_{i,k}$. In preliminary, the normal signals are computed in order to know the service set at each time instant k. Then, the attack input signal u_k is computed with the SLP to maximize the cost degradation function while keeping the service states $[x_5, x_6, x_7, x_8, x_9, x_{10}]$ inside X_s given for a $\delta = 10^{-4}$ error around the normal service. At the 60^{th} iteration of the SLP, the cost degradation function converges as shown in Figure. 7.11. It shows that the cost degradation function raises by 3.25 times from the normal signals.

7.4 Results analysis

For the AT1 on the film mfg. system, we can observe that the robust convex simulation-based method provides better results: the restrained input set $\hat{\mathcal{E}}_u$ is larger, allowing larger control actions. This observation might come from a tighter outer-approximation of the reachable set with the robust convex simulation-based method. In fact, it is less conservative than the invariant set based method as it allows the search for a symmetric matrix-valued polynomials $\tilde{Q}(t)$ rather than a symmetric matrix $\tilde{Q}$. However, this method consumes more computational resources.

By comparing the AT2 on both applications, we can notice that stealthy aging attacks have more capabilities on the satellite application, as the cost degradation function raises more. This observation might come from the physical system type: the satellite one comprises four motors, each of them acting on the three Euler angular rates, and then the Euler angles; whereas the film mfg. system comprises two motors, each of them acting on either $T1$ or $T2$, with T depending on $T1$ and $T2$. This analysis reveals that the satellite application is an over-actuated system. Note that it is over-actuated on purpose to tackle a potential (not intended) failure of a motor or reaction wheel. The stealthy aging attacks benefit from it by compensating the control action of one actuator with another one to deliver the desired service as shown in Figure. 7.11 (2^{nd} from top). However, for the not over-actuated type (film mfg. system), we can note that stealthy aging attacks benefit from the dynamic of the service to perform the attacks. In fact, the motor torque having a faster dynamic than the motor angular velocities, i.e. electrical and mechanical dynamic, it is possible to fastly change the control input $u(t)$ such that it is filtered by the motor angular velocity dynamic to remain stealthy.

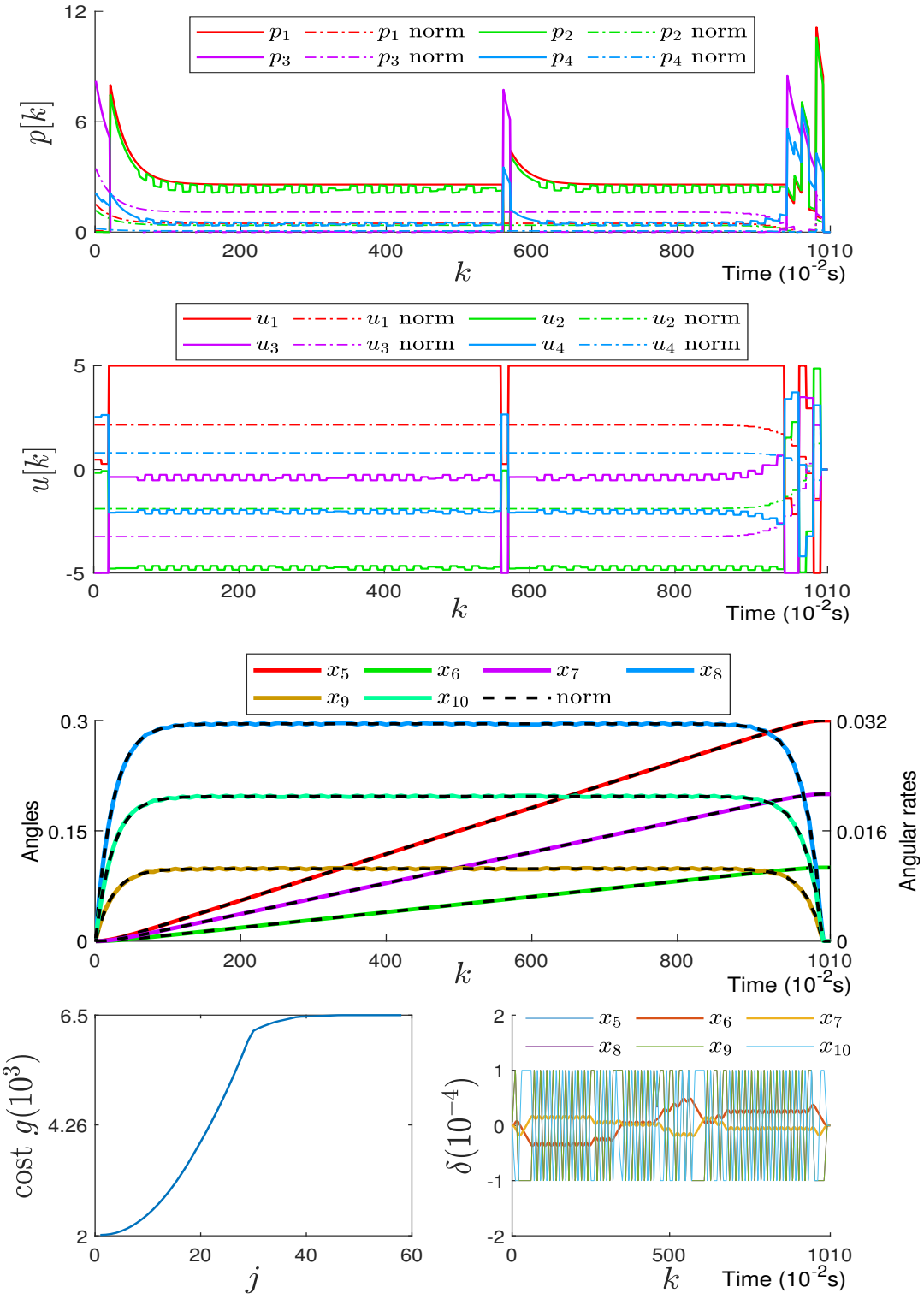


Figure 7.11: AT2 for the satellite: at the 60$^{\text{th}}$ iteration, the cost degradation function is maximized (1^{st} from bottom) with the degradation function (1^{st} from top) for the optimal attack input signal (2^{nd} from top) while delivering the desired service (2^{nd} from bottom)

7.5 Conclusion

In this Chapter, we have applied our methods upon two application examples. We have shown that the robut convex simulation-based method is more general and provides better results. However, the computational time is higher than the one with the invariant set-based method. In particular, we have observed that stealthy aging attacks seem more dangerous against dynamical systems exhibiting an over actuation as abnormal control actions can be compensated by the other actuators to hide the attacks regarding the delivered service.

Conclusion

Conclusion

The research works presented in this manuscript address the analysis and the prevention of stealthy aging attacks. Such attacks aim to accelerate the aging of the actuators, while the process continues to deliver the desired service. In particular, these research works considers an attacker that can manipulate the control input. Thus, their objective is to violate the process integrity.

The problem has been split into two cases of stealthy aging attacks: on one hand, a subcase of attacks (AT1) is considered. AT1 aims to let the dynamical system reach an abnormal aging set featuring an aging faster than normal. For AT1, the contribution of the methods resides in (i) proving the feasibility of such attacks, i.e. can they reach the process integrity? If they are feasible, (ii) restraining the control input set to make them unfeasible, i.e. the process integrity cannot be reached. On the other hand, the general case of stealthy aging attacks (AT2) is considered. AT2 aims to maximize the time integral of a degradation function, modeling the stresses acting on the dynamical system. The contribution of the method resides in (i) analyzing the potential aging an attacker could lead to the actuators.

On one hand, the methods for AT1 are based on the search of a pseudo-Lyapunov function for the dynamical system. This function defines κ-level set from which the state trajectory of the dynamical system will never go out. This allows the proof of the feasibility of such attacks. Two methods have been proposed: one based on the invariance concept, called the invariant set based method; and the other based on the robust simulation of the dynamical system, called the convex robust simulation-based method. The first method allows the analysis of stealthy aging attacks on an infinite-time horizon, while the second one allows it on a finite-time horizon. Thereafter, the methods search for a restrained control input set to make stealthy aging attacks unfeasible. Both methods solve a SDP problem with LMI constraints.

On the other hand, the method for AT2 is based on the construction of a control signal maximizing the time integral of a degradation function, modeling the stresses acting on the actuators. The idea is to find an optimal control signal under some constraints related to AT2. It allows the analysis of the impact stealthy aging attacks of AT2 could cause against a process. Since the degradation function is intrinsically nonlinear, then the method solves a SLP problem.

From a theoretical analysis of the methods, for AT1 the convex robust simulation-based one is more adapted to deal with stealthy aging attacks. Indeed, no condition on the Lyapunov stability is required for this method, unlike the invariant set-based one. Moreover, the invariant set-based-method has a computational drawback: the multiplier α needs to be fixed, but it is unbounded. As a result, one can search for an pseudo-Lyapunov function, but there is no guarantee that this is the optimal one. In addition,

the convex robust simulation-based method is less conservative. In fact, the method solves a Matrix SOS (MSOS) problem giving more degree of freedom. However, this method consumes more computational resources. About the method for AT2, it allows the analysis of potential impact of such attacks in the general case, but does not prove the feasibility of such attacks.

Lastly, the methods for AT1 and AT2 have been applied upon two applications: one from the manufacturing area, and another from the aerospace one. The results confirm the theoretical analysis: the convex robust simulation-based method provides a less restrictive condition on the control input set to guarantee the non-feasibility of AT1. About AT2, the results show that the optimal control signal maximizing the stresses can exploit (i) the dynamic property of the system to hide the attack, i.e. mechanical dynamic is slower than the electrical dynamic; (ii) the over-actuated property of the system to hide the attack.

This thesis has addressed the following scientific issues:

- Modeling the attack stealthiness in set theory,
- Stability condition required from the invariant set theory,
- Loss of the temporal variable in the invariant set theory.

From these research works, several research axes for pursuing these three years of research are proposed.
In the short term:

- From the aging of actuators, it seems relevant to map which stresses can occur from the control signals, and at which level it accelerates the aging of the actuator. In fact, the proposed methods are relevant only if one is able to describe the dangerous stresses, otherwise our methods cannot be applied. Furthermore, recalling that the stresses can be internal meaning that they are inherent in the operation of the actuator; or external meaning that they originate in the environment of the actuator. From this, it appears that both stresses need to be investigated. Indeed, an attacker might control another part of the process in order to increase the external stresses of an actuator.
- From the methods for AT1, the prevention of attacks consists in restraining the control input set. Such restriction has been formulated in terms of the control input set size, i.e. the trace of the shape matrix. However, the center has not been considered as a decision variable, otherwise the constraint becomes a BMI. By fixing the center, we find the largest restrained control input set for the fixed center, but this center is not guaranteed to give the largest set. Thus, this issue will be investigated.
- The methods for AT1 and AT2 will be applied on two testbeds to validate the methods. Two testbeds are available in our research lab: (i) a dc motor (drive)

interconnected with another dc motor (load) producing a load torque, and (ii) a drone with two dc motors. Those two testbeds are relevant as they share similar properties than the ones of the Applications studied in this thesis.

In the mid term:

- As a complement with the validation on a testbeds, it is common that the model does not fit with the real system owing to the modeling assumptions and the external disturbances acting on the process (input disturbance, output disturbance). So, the methods will be extended to a model with uncertainties.

- As seen in this thesis, the attacker requires to know the desired service, i.e. service set, in order to launch a stealthy aging attacks. We have stated in this thesis that the attacker can assess the expected service from the process model and the control signal transmitted by the controller. Thus, we will conduct research about the quality of knowledge the attacker has about the service set. Recalling that the detector in the controller side is based on a state estimator, can the attacker from the communication side compute the expected service around a small error compared to the one computed by the state estimator? This small error must guarantee the stealthiness of the attack.

- This thesis provides methods for analyzing and preventing stealthy aging attacks, but does not contribute to the detection of anomalies in the control signal. From the proposed methods, the convex robust simulation could be derived for detection purpose. In fact, it could be possible to place the detection system close to the process, i.e. after the transmission channel, and collect the control signal in real-time. Then, starting from an initial bounding state ellipsoid, the system checks if the bounding state ellipsoid at a future time will violate the process integrity. Obviously, it will require further conditions that might be strong such as (i) the control signal does not change during the computational time of the method, (ii) the computation time is guaranteed to be lower than the dynamic of the process.

- As presented in this thesis, the malevolence can be seen as a new root cause of failures in ICSs. It is intended and aims to reach the process integrity. However, the detection of an anomaly leads the detection system to assess the root cause. This leads to the problem of anomaly characterization. This is a crucial problem: if the malevolence is miscategorized and considered as another classical root cause by the detection system, then inappropriate actions will be take by the operators, without removing the malevolence. In a worst case, an attack could exploit it to induce failures at different period of times to be miscategorized in an intentional manner. We think that the characterization needs to focus on the specific characteristics of the malevolence. As stated in this thesis, the intentional characteristic combined with analysis of the propagation path could be an entry point to deal with. Notice that the propagation path of an anomaly needs a tracking of the anomaly which could be done by detecting anomalies in all the potential propagation paths as presented in this thesis. Obviously, this will lead further research works about

the effect in the anomaly characterization if some propagation paths taken by the anomaly are missing. About this, we think that the works in the confidence of the data based on Bayesian Network model could be appropriate.

In the long term:

- The methods for AT1 have a limit: they search for an ellipsoid that encompasses the state under some constraints. However, ellipsoid might not be the appropriate set for being the smallest set. Other sets exist as mentioned in this thesis. Further research works will develop methods with the same objective, but with other sets. These research works will develop a method searching for various sets, including the ellipsoids presented in this thesis. By definition all the sets we will find contain the state trajectory, thus the set resulting from the intersection of all the sets contains also the state trajectory. A tighter outer-approximation of the state set could result from this approach.

- As seen in this thesis, with the methods for AT1 one can prove the feasibility of stealthy aging attacks and prevent them to occur by computing the largest restrained control input set. However, it has not been extended to the general case AT2. Future research works will investigate it.

- In future works we will extend our methods to the class of PWA system. In fact, it allows a description of a physical system that follows an affine time-invariant model for some regions of the state. This extension is important as it allows, among others, the modeling of the physical entity that deliver the energy to the actuator, called the pre-actuator (e.g. power converter, pneumatic distributor). In fact, their model is often non-linear.

- The design of any system (e.g. controller, detector) requires a model of the process. This model is obtained under some modeling assumptions to simplify the problem. Those modeling assumptions are validated in particular cases. However, no system checks if those modeling assumptions are satisfied in realtime. So, our future research work will concern it around two main questions: can an attacker exploit the modeling assumptions to launch stealthy aging attacks? and how can we design detection system to verify in realtime that those modeling assumptions are satisfied?

Appendices

Appendix A

Matlab program: Invariant set-based method

```
%% Step 1 and Step 2
At = [A a; zeros(1,n) 0];
Bt = [B; zeros(1,m)];
Gam = [zeros(1,n) 1];
Xit = [Xi zeros(n,1); zeros(1,n) 0];
xit = [xib; 0];
Qt = sdpvar(n+1);
beta = sdpvar(1,1);
lambda = sdpvar(1,1);
nothing = sdpvar(1);
con=(Qt(1:n,1:n)-eps≥0);
con=con+(beta≥0);
con=con+(lambda≥0);
con=con+(nothing≥0);
ellSize = 0;
 for alpha=alphaMin:DeltaAlpha:alphaMax
    M = [At'*Qt+Qt*At Qt*Bt;
         Bt'*Qt zeros(m,m)];

    N = [Qt-Gam'*Gam zeros(n+1,m);
         zeros(m, n+1) zeros(m,m)];

    S = [Gam'*Gam-Gam'*ub'*R*ub*Gam Gam'*ub'*R;
         R*ub*Gam -R];

    T = [Gam'*Gam-Xit+Xit*xit*Gam+Gam'*xit'*Xit-Gam'*xit'*Xit*xit*Gam ...
        zeros(n+1,m);
         zeros(m,n+1) zeros(m,m)];

    con(end)=([-M-alpha*N-beta*S-lambda*T]≥0);
    % Solve the optimization problem OP_1
    diag=optimize(con,-trace(Qt),sdpsettings('solver','mosek','verbose',0,
                                             'cachesolvers',1));
     if (diag.problem==0 || diag.problem==4) && (all(eig(value(Qt))≥0))
        Qtval = value(Qt);
        % If the ellipsoid size is lower than the stored one ...
        if trace(value(Qt))> ellSize
            % ... Then store the ellipsoid matrix
            Qt_stck=Qtval;
            ellSize = trace(Qtval);
        end
    end
```

```
 end
% Extract the ellipsoid matrix and the center
Qn=Qt_stck;
xn_out = -Qn([1:n],[1:n])\Qn([1:n],n+1);
% Normalization
Qn_out = Qn([1:n],[1:n])/(1-Qn(n+1,n+1)+xn_out'*Qn([1:n],[1:n])*xn_out);
```

```
%% Step 3
At = [A a; zeros(1,n) 0];
Bt = [B; zeros(1,m)];
Gam = [zeros(1,n) 1];
Xit = [Xi zeros(n,1); zeros(1,n) 0];
xit = [xib; 0];
% Parameters of the bisection method
R_ov = Δ_ov*R0; %Should not hit
R_und = Δ_und*R0; %Should hit
% Start the bisection method
Δ_wrk = Δ_ov;
while(Δ_ov-Δ_und > tol)
    Δ = (Δ_ov+Δ_und)/2
    R = Δ*R0; % Value of R at the first iteration
    bool = 0; %0 if no hit, otherwise 1
    eps = 10^-8;
    Qt = sdpvar(n+1);
    beta = sdpvar(1,1);
    lambda = sdpvar(1,1);
    nothing = sdpvar(1);
    con=(Qt(1:n,1:n)-eps≥0);
    con=con+(beta≥0);
    con=con+(lambda≥0);
    con=con+(nothing≥0);
    ellSize = 0;
    for alpha=alphaMin:DeltaAlpha:alphaMax
        M = [At'*Qt+Qt*At Qt*Bt;
             Bt'*Qt zeros(m,m)];

        N = [Qt-Gam'*Gam zeros(n+1,m);
            zeros(m, n+1) zeros(m,m)];

        S = [Gam'*Gam-Gam'*ub'*R*ub*Gam Gam'*ub'*R;
            R*ub*Gam -R];

        T = ...
            [Gam'*Gam-Xit+Xit*xit*Gam+Gam'*xit'*Xit-Gam'*xit'*Xit*xit*Gam ...
            zeros(n+1,m);
            zeros(m,n+1) zeros(m,m)];

        con(end)=([-M-alpha*N-beta*S-lambda*T]≥0);
        % Solve the optimization problem OP_1
        diag=optimize(con,-trace(Qt),sdpsettings('solver','mosek',
                                    'verbose',0,'cachesolvers',1));
        if (diag.problem==0 || diag.problem==4) && (all(eig(value(Qt))≥0))
            Qtval = value(Qt);
```

```
45              % If the ellipsoid size is lower than the stored one ...
46              if trace(value(Qt))> ellSize
47                  % ... Then store the ellipsoid matrix
48                  Qt_stck=Qtval;
49                  ellSize = trace(Qtval);
50              end
51          end
52      end
53      % Extract the ellipsoid matrix and the center
54      Qn=Qt_stck;
55      xn_out = -Qn([1:n],[1:n])\Qn([1:n],n+1);
56      % Normalization
57      Qn_out = ...
            Qn([1:n],[1:n])/(1-Qn(n+1,n+1)+xn_out'*Qn([1:n],[1:n])*xn_out);
58      % Verify if the ellipsoid hits with the abnormal aging set (2
59      % halfspaces in this case)
60      check1 = (sqrt(cHyp1'*(Qn_out\cHyp1))-abs(mu-cHyp1'*xn_out))/
61                                                  (sqrt(cHyp1'*cHyp1));
62      check2 = (sqrt(cHyp2'*(Qn_out\cHyp2))-abs(mu-cHyp2'*xn_out))/
63                                                  (sqrt(cHyp2'*cHyp2));
64      % If it hits ...
65      if (check1>0 || check2>0)
66          % ... Then reduce the size of R
67          Δ_und = Δ;
68      else
69          % ... Otherwise increase the size of R
70          Δ_ov = Δ;
71          Δ_wrk = Δ;
72          % Store the ellipsoid matrix and the center
73          Qn_out_wrk = Qn_out
74          xn_out_wrk = xn_out
75      end
76         % Clear memory
77         yalmip('clear');
78         % Define decision variables for the next iteration step
79          nothing = sdpvar(1);
80  end
81  % Output: state and input ellipsoid matrices and centers
82  Qn_out_opt = Qn_out_wrk;
83  xn_out_opt = xn_out_wrk;
84  R = Δ_wrk * R0;
85  ub
```

Appendix B

Matlab program: Robust convex simulation-based method

```
%% Step 1 and Step 2
Atil = [A, ab; zeros(1,n), 0];
Btil = [B; zeros(1,m)];
Gam = [zeros(1,n) 1];
Xit = [Xi zeros(n,1); zeros(1,n) 0];
xit = [xib; 0];
Qsi_org = Xi;
qsi_org = xib;
Qsitil = [Qsi_org, zeros(n,1);
          zeros(1,n) 0];
qsitil = [qsi_org; 0];
Itil = eye(n);

% Initial condition Eps_x(E0,c0)
E0 = Eini;
c0 = cini;

% Time parameters
tf = 0.5; % Total simulation time [s]
Delta_t = 0.001; % Time step [s]
N = tf/(Delta_t); % Number of iterations

% Initialization
E([1:n],[1:n]) = E0;
c([1:n],1) = c0;

% To store the hit
ck=zeros(1,round(N));

% Verify if at initial time it hits the abnormal aging set
check1 = (sqrt(cHyp1'*(E0\cHyp1))-abs(mu-cHyp1'*c0))/(sqrt(cHyp1'*cHyp1));
check2 = (sqrt(cHyp2'*(E0\cHyp2))-abs(mu-cHyp2'*c0))/(sqrt(cHyp2'*cHyp2));
if(check1>0 || check2>0)
    ck(1) = 1;
    disp('error')
    pause();
else
    ck(1)=0;
end

% Time SOS degree 0
t=sdpvar(1);
```

```
% beta SOS degree 0
beta = sdpvar(1);
% Alpha SOS degree 6
a0=sdpvar(1);
a1=sdpvar(1);
a2=sdpvar(1);
a3=sdpvar(1);
a4=sdpvar(1);
a5=sdpvar(1);
a6=sdpvar(1);
% Q MSOS degree 6
Q1=sdpvar(n+1);
Q2=sdpvar(n+1);
Q3=sdpvar(n+1);
Q4=sdpvar(n+1);
Q5=sdpvar(n+1);
Q6=sdpvar(n+1);
% Z MSOS degree 0
Z=sdpvar(n+m+1);
% X MSOS degree 0
X = sdpvar(n);
nothing = sdpvar(1);
% Constraints
con=(Z≥0);
con = con +(X≥0);
con = con + (beta≥0);
con=con+(nothing≥0);
gam = t*(tf-t);
j = 1;

for i=1:1:N
    Q0 = [E0    -E0*c0;
          -c0'*E0  c0'*E0*c0];

    % alpha(t)
    alpha = a0+a1*t+a2*t^2+a3*t^3+a4*t^4+a5*t^5+a6*t^6;
    % Q(t)
    Q=Q0+Q1*t+Q2*t^2+Q3*t^3+Q4*t^4+Q5*t^5+Q6*t^6;
    % Qdot(t)
    Qdot=Q1+2*Q2*t+3*Q3*t^2+4*Q4*t^3+5*Q5*t^4+6*Q6*t^5;
    % Objective: Q(tf)
    Qf=Q0+Q1*tf+Q2*tf^2+Q3*tf^3+Q4*tf^4+Q5*tf^5+Q6*tf^6;
    con(end)=(Qf(1:n,1:n)-epss≥0);

    M = [Qdot + Atil'*Q + Q*Atil, Q*Btil;
         (Q*Btil)', zeros(m)];

    S = [Gam'*Gam - Gam'*ub'*R*ub*Gam, Gam'*ub'*R;
         (Gam'*ub'*R)', -R];

    T = [Gam'*Gam - Qsitil + Qsitil*qsitil*Gam + Gam'*qsitil'*Qsitil - ...
        Gam'*qsitil'*Qsitil*qsitil*Gam, zeros(n+1,m);
         zeros(m,n+1), zeros(m,m)];

    % Solve MSOS problem
```

```
97      diag = solvesos([con sos(-M-alpha*S-beta*T-gam*Z) ...
            sos(Q([1:n],[1:n])-gam*X - epsa*Itil) ...
            sos(alpha)],-trace(Qf(1:n+1,1:n+1)),sdpsettings('solver','mosek',
98              'verbose',0,'cachesolvers',1),
99              [beta;a0;a1;a2;a3;a4;a5;a6;Q1(:);Q2(:);Q3(:);Q4(:);Q5(:);Q6(:);
100             Z(:);X(:)]);
101     if(diag.problem==1)
102         pause();
103     end
104     % Extract the ellipsoid matrix and the center
105     QfVal = value(Qf);
106     xb_out = -QfVal([1:n],[1:n])\QfVal([1:n],n+1);
107     % Normalization
108     Q_out = QfVal([1:n],[1:n])/(1-QfVal(n+1,n+1)+
109                 xb_out'*QfVal([1:n],[1:n])*xb_out);
110     % Store the output at iteration i
111     if(mod(i,loop_nb) == 0)
112         E_stock([(j+(j-1)*(n-1)):(j*n)],[1:n]) = Q_out;
113         c_stock([1:n],j) = xb_out;
114         j = j+1
115     end
116     % Verify if it hits with the abnormal aging set
117     check1 = (sqrt(cHyp1'*(Q_out\cHyp1))-abs(mu-cHyp1'*xb_out))/
118                                                 (sqrt(cHyp1'*cHyp1));
119     check2 = (sqrt(cHyp2'*(Q_out\cHyp2))-abs(mu-cHyp2'*xb_out))/
120                                                 (sqrt(cHyp2'*cHyp2));
121     if(check1>0 || check2>0)
122         % It hits
123         ck(i+1) = 1;
124     else
125         % It does not hit
126         ck(i+1)=0;
127     end
128
129     % Clear memory
130     if(mod(i,25) == 0)
131        % Clear memory
132        yalmip('clear');
133        % Define decision variables for the next iteration step
134        nothing = sdpvar(1);
135         % SOS degree 0
136         t=sdpvar(1);
137         % beta SOS degree 0
138         beta=sdpvar(1);
139         % Alpha SOS degree 6
140         a0=sdpvar(1);
141         a1=sdpvar(1);
142         a2=sdpvar(1);
143         a3=sdpvar(1);
144         a4=sdpvar(1);
145         a5=sdpvar(1);
146         a6=sdpvar(1);
147         % Q MSOS degree 6
148         Q1=sdpvar(n+1);
149         Q2=sdpvar(n+1);
```

```
150          Q3=sdpvar(n+1);
151          Q4=sdpvar(n+1);
152          Q5=sdpvar(n+1);
153          Q6=sdpvar(n+1);
154          % Z MSOS degree 0
155          Z=sdpvar(n+m+1);
156          % X MSOS degree 0
157          X = sdpvar(n);
158          % Constraints
159          con=(Z≥0);
160          con = con +(X≥0);
161          con = con + (beta≥0);
162          con=con+(nothing≥0);
163          gam = t*(tf-t);
164      end
165      % Final ellipsoid at iteration i is then the initial ellipsoid at i+1
166      E0 = Q_out;
167      c0 = xb_out;
168  end
```

```
1  %% Step 3
2  Atil = [A, ab; zeros(1,n), 0];
3  Btil = [B; zeros(1,m)];
4  Gam = [zeros(1,n) 1];
5  Xit = [Xi zeros(n,1); zeros(1,n) 0];
6  xit = [xib; 0];
7  Qsi_org = Xi;
8  qsi_org = xib;
9  Qsitil = [Qsi_org, zeros(n,1);
10           zeros(1,n) 0];
11 qsitil = [qsi_org; 0];
12 Itil = eye(n);
13
14 % Initial condition Eps_x(E0,c0)
15 E0 = Eini;
16 c0 = cini;
17
18 % Time parameters
19 tf = 0.5; % Total simulation time [s]
20 Delta_t = 0.001; % Time step [s]
21 N = tf/(Delta_t); % Number of iterations
22
23 % Initialization
24 E([1:n],[1:n]) = E0;
25 c([1:n],1) = c0;
26
27 % To store the hit
28 ck=zeros(1,round(N));
29
30 % Verify if at initial time it hits the abnormal aging set
31 check1 = (sqrt(cHyp1'*(E0\cHyp1))-abs(mu-cHyp1'*c0))/(sqrt(cHyp1'*cHyp1));
32 check2 = (sqrt(cHyp2'*(E0\cHyp2))-abs(mu-cHyp2'*c0))/(sqrt(cHyp2'*cHyp2));
33 if(check1>0 || check2>0)
```

```
34      ck(1) = 1;
35      disp('error')
36      pause();
37  else
38      ck(1)=0;
39  end
40
41  % Time SOS degree 0
42  t=sdpvar(1);
43  % beta SOS degree 0
44  beta = sdpvar(1);
45  % Alpha SOS degree 6
46  a0=sdpvar(1);
47  a1=sdpvar(1);
48  a2=sdpvar(1);
49  a3=sdpvar(1);
50  a4=sdpvar(1);
51  a5=sdpvar(1);
52  a6=sdpvar(1);
53  % Q MSOS degree 6
54  Q1=sdpvar(n+1);
55  Q2=sdpvar(n+1);
56  Q3=sdpvar(n+1);
57  Q4=sdpvar(n+1);
58  Q5=sdpvar(n+1);
59  Q6=sdpvar(n+1);
60  % Z MSOS degree 0
61  Z=sdpvar(n+m+1);
62  % X MSOS degree 0
63  X = sdpvar(n);
64  nothing = sdpvar(1);
65  % Constraints
66  con=(Z≥0);
67  con = con +(X≥0);
68  con = con + (beta≥0);
69  con=con+(nothing≥0);
70  gam = t*(tf-t);
71
72  % Parameters of the bisection method
73  R_ov = Δ_ov*R0; %Should not hit
74  R_und = Δ_und*R0; %Should hit
75
76  % Start the bisection method
77  Δ_wrk = Δ_ov;
78  while(Δ_ov-Δ_und > tol)
79      Δ = (Δ_ov+Δ_und)/2
80      R = Δ*R0;
81      i = 1;
82      j = 1;
83      bool = 0; %0 if no hit, otherwise 1
84      E0 = Eini;
85      c0 = cini;
86
87      % Compute the evolution of state ellipsoids for the N iterations while
88      % none ellipsoid hits with the abnormal aging set
```

```
89      while (i≤N && bool==0)
90          Q0 = [E0    -E0*c0;
91                -c0'*E0  c0'*E0*c0];
92
93          % alpha(t)
94          alpha = a0+a1*t+a2*t^2+a3*t^3+a4*t^4+a5*t^5+a6*t^6;
95          % Q(t)
96          Q=Q0+Q1*t+Q2*t^2+Q3*t^3+Q4*t^4+Q5*t^5+Q6*t^6;
97          % Qdot(t)
98          Qdot=Q1+2*Q2*t+3*Q3*t^2+4*Q4*t^3+5*Q5*t^4+6*Q6*t^5;
99          % Objective: Q(tf)
100         Qf=Q0+Q1*tf+Q2*tf^2+Q3*tf^3+Q4*tf^4+Q5*tf^5+Q6*tf^6;
101         con(end)=(Qf(1:n,1:n)-epss≥0);
102
103         M = [Qdot + Atil'*Q + Q*Atil, Q*Btil;
104              (Q*Btil)', zeros(m)];
105
106         S = [Gam'*Gam - Gam'*ub'*R*ub*Gam, Gam'*ub'*R;
107              (Gam'*ub'*R)', -R];
108
109         T = [Gam'*Gam - Qsitil + Qsitil*qsitil*Gam + ...
                 Gam'*qsitil'*Qsitil - Gam'*qsitil'*Qsitil*qsitil*Gam, ...
                 zeros(n+1,m);
110              zeros(m,n+1), zeros(m,m)];
111
112         % Solve MSOS problem
113         diag = solvesos([con sos(-M-alpha*S-beta*T-gam*Z) ...
                 sos(Q([1:n],[1:n])-gam*X - epsa*Itil) ...
                 sos(alpha)],-trace(Qf(1:n+1,1:n+1)),sdpsettings('solver',
114         'mosek','verbose',0,'cachesolvers',1),
115         [beta;a0;a1;a2;a3;a4;a5;a6;Q1(:);Q2(:);Q3(:);Q4(:);Q5(:);Q6(:);
116         Z(:);X(:)]);
117         if(diag.problem==1)
118             pause();
119         end
120         % Extract the ellipsoid matrix and the center
121         QfVal = value(Qf);
122         xb_out = -QfVal([1:n],[1:n])\QfVal([1:n],n+1);
123         % Normalization
124         Q_out = QfVal([1:n],[1:n])/(1-QfVal(n+1,n+1)+xb_out'*
125         QfVal([1:n],[1:n])*xb_out);
126         % Store the output at iteration i
127         if(mod(i,loop_nb) == 0)
128             E_stock([(j+(j-1)*(n-1)):(j*n)],[1:n]) = Q_out;
129             c_stock([1:n],j) = xb_out;
130             j = j+1;
131         end
132         % Verify if it hits with the abnormal aging set
133         check1 = (sqrt(cHyp1'*(E0\cHyp1))-abs(mu-cHyp1'*c0))/
134                                          (sqrt(cHyp1'*cHyp1));
135         check2 = (sqrt(cHyp2'*(E0\cHyp2))-abs(mu-cHyp2'*c0))/
136                                          (sqrt(cHyp2'*cHyp2));
137
138         if(check1>0 || check2>0)
139             % It hits
```

```
140             ck(i+1) = 1;
141             bool=1;
142         else
143             % It does not hit
144             ck(i+1)=0;
145         end
146 
147         % Clear memory
148         if(mod(i,25) == 0)
149             % Clear memory
150             yalmip('clear');
151             % Define decision variables for the next iteration step
152             nothing = sdpvar(1);
153             % SOS degree 0
154             t=sdpvar(1);
155             % beta SOS degree 0
156             beta=sdpvar(1);
157             % Alpha SOS degree 6
158             a0=sdpvar(1);
159             a1=sdpvar(1);
160             a2=sdpvar(1);
161             a3=sdpvar(1);
162             a4=sdpvar(1);
163             a5=sdpvar(1);
164             a6=sdpvar(1);
165             % Q MSOS degree 6
166             Q1=sdpvar(n+1);
167             Q2=sdpvar(n+1);
168             Q3=sdpvar(n+1);
169             Q4=sdpvar(n+1);
170             Q5=sdpvar(n+1);
171             Q6=sdpvar(n+1);
172             % Z MSOS degree 0
173             Z=sdpvar(n+m+1);
174             % X MSOS degree 0
175             X = sdpvar(n);
176             % Constraints
177             con=(Z≥0);
178             con = con +(X≥0);
179             con = con + (beta≥0);
180             con=con+(nothing≥0);
181             gam = t*(tf-t);
182         end
183         % Final ellipsoid at iteration i is then the initial ellipsoid ...
                at i+1
184         E0 = Q_out;
185         c0 = xb_out;
186         i = i+1;
187     end
188     % If it hits ...
189     if (check1>0 || check2>0)
190         % ... Then reduce the size of R
191         Δ_und = Δ;
192     else
193         % ... Otherwise increase the size of R
```

```
194          Δ_ov = Δ;
195          Δ_wrk = Δ;
196          % Store the evolution of state ellipsoids
197          Q_out_wrk = Q_out;
198          xb_out_wrk = xb_out
199          E_out_wrk(:,:) = E_stock(:,:);
200          c_out_wrk(:,:) = c_stock(:,:);
201      end
202      % Clear memory
203      yalmip('clear');
204      % Define decision variables for the next iteration step
205      nothing = sdpvar(1);
206      % SOS degree 0
207      t=sdpvar(1);
208      % beta SOS degree 0
209      beta=sdpvar(1);
210      % Alpha SOS degree 6
211      a0=sdpvar(1);
212      a1=sdpvar(1);
213      a2=sdpvar(1);
214      a3=sdpvar(1);
215      a4=sdpvar(1);
216      a5=sdpvar(1);
217      a6=sdpvar(1);
218      % Q MSOS degree 6
219      Q1=sdpvar(n+1);
220      Q2=sdpvar(n+1);
221      Q3=sdpvar(n+1);
222      Q4=sdpvar(n+1);
223      Q5=sdpvar(n+1);
224      Q6=sdpvar(n+1);
225      % Z MSOS degree 0
226      Z=sdpvar(n+m+1);
227      % X MSOS degree 0
228      X = sdpvar(n);
229      % Constraints
230      con=(Z≥0);
231      con = con +(X≥0);
232      con = con + (beta≥0);
233      con=con+(nothing≥0);
234      gam = t*(tf-t);
235  end
236  % Output: state and input ellipsoid matrices and centers
237  E_stock = E_out_wrk;
238  c_stock = c_out_wrk;
239  R = Δ_wrk*R0;
240  ub
```

Appendix C

Matlab program: Optimal attack input signal

```
%% Step 1 : Discretization
At = [A a; zeros(1,n) 0];
Bt = [B; zeros(1,m)];
sys = ss(At,Bt,C,D);
% Discretization
Te = 0.001; % Sampling time
opt = c2dOptions('Method','tustin');
sysd = c2d(sys,Te,opt);
% Transformation for having Cd = eye(ns)
T = inv(eye(n+1))*sysd.C;
sysd = ss2ss(sysd,T);
At = sysd.A;
Bt = sysd.B;
C = sysd.C;
D = sysd.D;
save model.mat At Bt C D Te n m out
```

```
%% Step 2 : Parametrization
clear all;
load model_presentation.mat;

N=2000; % Total number of time instants
gtemp = 50; % Number of time instants stored in a matrix (reduces ...
    computational time)
Nb=N/gtemp;
F=sparse(N*(n+1),n+1);
for i=1:N
    F([(i-1)*(n+1)+1:i*(n+1)],:)=At^i;
end
temp = sparse(n+1,N*m);
i=0;
while i<N
    i=i+1;
    temp(:,[m*(i-1)+1:m*i])=At^(N-i)*Bt;
end
for j=1:Nb
    j
    Gtemp = sparse(gtemp*(n+1),N*m);
    for i=1:gtemp
        Gtemp([(i-1)*(n+1)+1:i*(n+1)],[1:m*(i+(j-1)*gtemp)])=
```

```
                        temp(:,[end-m*(i+(j-1)*gtemp)+1:end]);
    end
    if j≠1
        G = [G; Gtemp];
    else
        G=Gtemp;
    end
end

save param_model_presentation.mat F G N -v7.3 % Added
```

```
%% Step 3 : Verify if the normal trajectories satisfy the constraints
yalmip('clear');
clear all;
close all;
load model_presentation.mat;
load param_model_presentation.mat;

H = sparse(N*(n+1),N*(n+1));
I = sparse(N*(n+1),N*m);
for i =1:N
    for j=1:n+1
        H(j+(n+1)*(i-1),j+(n+1)*(i-1)) = 1;
    end
    I([1+(n+1)*(i-1):i*(n+1)],[1+m*(i-1):i*m]) = D;
end

x=sdpvar((n+1)*N,1);
y = sdpvar((n+1)*N,1);
u=sdpvar(m*N,1);
xi = [zeros(n,1);1];

con=(x(:)==F*xi+G*u);
con=con+(y(:)==H*[xi;x(1:end-(n+1))]+I*u);

for i=1:(n+1)*N
    if mod(i,3)==0
       x(i)=1;
    end
end

con=con+(u(:)≤uh);
con=con+(u(:)≥ul);

disp('In progress')

obj=0;
u = out.control_normal(2:end,2);
j=2;
k=2;
for i=1:(n+1)*N
    if mod(i,3)==1
        y(i) = out.velocity_normal(j,2);
```

```
        j=j+1;
    elseif mod(i,3)==2
        y(i) = out.current_normal(k,2);
        k=k+1;
    elseif mod(i,3)==0
        y(i) = 1;
    end
end
j=optimize(con,obj,sdpsettings('solver','mosek','verbose',1,
                                                'cachesolvers',1));
uopt=value(u);
xopt=value(x);
yopt=value(y);
save sol_model_presentation.mat xopt yopt uopt xi H I

% Extract value
X=zeros(n+1,N);
Y=zeros(n+1,N);
U=zeros(m,N);
for i=1:N
   X(:,i) = xopt((i-1)*(n+1)+1:i*(n+1));
   Y(:,i) = yopt((i-1)*(n+1)+1:i*(n+1));
   U(:,i) = uopt((i-1)*m+1:i*m);
end
```

```
%% Step 4 : Compute a feasible attack
yalmip('clear');
clear all;
close all;
load model_presentation.mat;
load param_model_presentation.mat;
load sol_model_presentation.mat;

nbServ = 1;
x=sdpvar((n+1)*N,1);
y = sdpvar((n+1)*N,1);
u=sdpvar(m*N,1);
ybnd = sdpvar(nbServ*N,1);
bnd_h=sparse(nbServ*N,1);
bnd_l=sparse(nbServ*N,1);
diff = 1e-1; % Trajectory error
bndFix=[diff NaN]';

con=(x(:)==F*xi+G*u);
con=con+(y(:)==H*[xi;x(1:end-(n+1))]+I*u);

for i=1:(n+1)*N
    if mod(i,3)==0
       x(i)=1;
    end
end
j=1;
for i =1:N*(n+1)
```

```
29      if(mod(i,3)==1)
30          ybnd(j,1)=x(i);
31          bnd_h(j,1)=xopt(i)+bndFix(1);
32          bnd_l(j,1)=xopt(i)-bndFix(1);
33          j=j+1;
34      end
35  end
36  con=con+(ybnd≤bnd_h);
37  con=con+(ybnd≥bnd_l);
38  M = 1; % Time step instant between each manipulation of the control signal
39  if(M>1)
40  temp=ceil(N/M)-1;
41  conAttck=sdpvar(temp*(M-1)*m+((N-1)-temp*M+1)*m,1);
42  ucon=sdpvar(temp*m+(N-1)-temp*M,1);
43
44  for i=1:temp
45      for j=1:M-1
46      conAttck(m*(i-1)*(M-1)+(j-1)*m+1:m*(i-1)*
47      (M-1)+(j-1)*m+m,1)=u(m*(i-1)*M+1:m*(i-1)*M+m);
48      ucon(m*(i-1)*(M-1)+(j-1)*m+1:m*(i-1)*
49      (M-1)+(j-1)*m+m,1)=u((i-1)*M*m+j*m+1:(i-1)*M*m+j*m+m);
50      end
51  end
52  j=1;
53  for i=temp*M:N-1
54      conAttck(m*(temp-1)*(M-1)+(M-1-1)*m+m+(j-1)*m+1:m*(temp-1)*
55      (M-1)+(M-1-1)*m+m+j*m)=u(temp*M*m+1:temp*M*m+m);
56      ucon(m*(temp-1)*(M-1)+(M-1-1)*m+m+(j-1)*m+1:m*(temp-1)*
57      (M-1)+(M-1-1)*m+m+j*m)=u(i*m+1:i*m+m);
58      j=j+1;
59  end
60      con=con+(ucon==conAttck);
61  end
62  con=con+(u(:)≤uh);
63  con=con+(u(:)≥ul);
64  obj=u'*u; % Whatever objective cost because the objective is to find ...
        feasible attack signals
65
66  j=optimize(con,obj,sdpsettings('solver','mosek','verbose',1,
67                                          'cachesolvers',1));
68  ufeas=value(u);
69  xfeas=value(x);
70  yfeas=value(y);
71  objfeas = value(obj);
72  save sol_pre_presentation.mat xfeas yfeas ufeas objfeas M bndFix
73
74  % Extract value
75  X=zeros(n+1,N);
76  Y=zeros(n+1,N);
77  U=zeros(m,N);
78  for i=1:N
79     X(:,i) = xfeas((i-1)*(n+1)+1:i*(n+1));
80     Y(:,i) = yfeas((i-1)*(n+1)+1:i*(n+1));
81     U(:,i) = ufeas((i-1)*m+1:i*m);
82  end
```

```
%% Step 5 : Compute the optimal attack signals
yalmip('clear');
clear all;
close all;
load model_presentation.mat;
load param_model_presentation.mat;
load sol_model_presentation.mat;
load sol_pre_presentation.mat;

nbServ = 1;
x=sdpvar((n+1)*N,1);
y=sdpvar((n+1)*N,1);
u=sdpvar(m*N,1);
xbnd = sdpvar(nbServ*N,1);
bnd_h=sparse(nbServ*N,1);
bnd_l=sparse(nbServ*N,1);
con=(x(:)==F*xi+G*u);
con=con+(y(:)==H*[xi;x(1:end-(n+1))]+I*u);

for i=1:(n+1)*N
    if mod(i,3)==0
       x(i)=1;
    end
end

vecA = sdpvar(N1,1);
for i=1:N1
    vecA(i) = y(end-(n+1)*i+1);
end
con=con+(vecA==ones(N1,1)*yopt(end-(n+1)*i+1);

j=1;
for i =1:N*n
    if(mod(i,2)==1)
        ybnd(j,1)=y(i);
        bnd_h(j,1)=yopt(i)+bndFix(1);
        bnd_l(j,1)=yopt(i)-bndFix(1);
        j=j+1;
    end
end
con=con+(ybnd≤bnd_h);
con=con+(ybnd≥bnd_l);

M=1;
if(M>1)
temp=ceil(N/M)-1;
conAttck=sdpvar(temp*(M-1)*m+((N-1)-temp*M+1)*m,1);
ucon=sdpvar(temp*m+(N-1)-temp*M,1);

for i=1:temp
    for j=1:M-1
    conAttck(m*(i-1)*(M-1)+(j-1)*m+1:m*(i-1)*
    (M-1)+(j-1)*m+m,1)=u(m*(i-1)*M+1:m*(i-1)*M+m);
    ucon(m*(i-1)*(M-1)+(j-1)*m+1:m*(i-1)*
```

```
55      (M-1)+(j-1)*m+m,1)=u((i-1)*M*m+j*m+1:(i-1)*M*m+j*m+m);
56      end
57  end
58  j=1;
59  for i=temp*M:N-1
60      conAttck(m*(temp-1)*(M-1)+(M-1-1)*m+m+(j-1)*m+1:m*(temp-1)*
61      (M-1)+(M-1-1)*m+m+j*m)=u(temp*M*m+1:temp*M*m+m);
62      ucon(m*(temp-1)*(M-1)+(M-1-1)*m+m+(j-1)*m+1:m*(temp-1)*
63      (M-1)+(M-1-1)*m+m+j*m)=u(i*m+1:i*m+m);
64      j=j+1;
65  end
66      con=con+(ucon==conAttck);
67  end
68
69  for it=1:200
70      % Evolution of uit constraint to have a linearized cost function close
71      % to the true cost function
72       con(end-1)=(u≤uit+0.1);
73       con(end)=(u≥uit-0.1);
74      % Objective function
75      for i=1:N
76      currentit(i) = xit(1+(n+1)*(i-1));
77      current(i) = x(1+(n+1)*(i-1));
78      end
79      objit=sum(currentit.^2);
80      objEvol(it) = objit;
81      obj=objit+sum(2*(current-currentit).*currentit);
82      j=optimize(con,-obj,sdpsettings('solver','mosek','verbose',1,
83                                                  'cachesolvers',1));
84      if(j.problem≠0)
85          disp('Value of diag:');
86          j
87          pause();
88      end
89      yit=value(y);
90      uit=value(u);
91
92      % Initialize
93      if(mod(jinst,25) == 0)
94         % Clear memory
95          yalmip('clear');
96          x=sdpvar((n+1)*N,1);
97          y=sdpvar((n+1)*N,1);
98          u=sdpvar(m*N,1);
99          xbnd = sdpvar(nbServ*N,1);
100         bnd_h=sparse(nbServ*N,1);
101         bnd_l=sparse(nbServ*N,1);
102         con=(x(:)==F*xi+G*u);
103         con=con+(y(:)==H*[xi;x(1:end-(n+1))]+I*u);
104         for i=1:(n+1)*N
105             if mod(i,3)==0
106                x(i)=1;
107             end
108         end
109
```

```
110          vecA = sdpvar(N1,1);
111          for i=1:N1
112              vecA(i) = y(end-(n+1)*i+1);
113          end
114          con=con+(vecA==ones(N1,1)*yopt(end-(n+1)*i+1);
115         j=1;
116          for i =1:N*n
117              if(mod(i,2)==1)
118                  ybnd(j,1)=y(i);
119                  bnd_h(j,1)=yopt(i)+bndFix(1);
120                  bnd_l(j,1)=yopt(i)-bndFix(1);
121                  j=j+1;
122              end
123          end
124          con=con+(ybnd≤bnd_h);
125          con=con+(ybnd≥bnd_l);
126
127          if(M>1)
128          temp=ceil(N/M)-1;
129          conAttck=sdpvar(temp*(M-1)*m+((N-1)-temp*M+1)*m,1);
130          ucon=sdpvar(temp*m+(N-1)-temp*M,1);
131
132          for i=1:temp
133              for j=1:M-1
134              conAttck(m*(i-1)*(M-1)+(j-1)*m+1:m*(i-1)*
135              (M-1)+(j-1)*m+m,1)=u(m*(i-1)*M+1:m*(i-1)*M+m);
136              ucon(m*(i-1)*(M-1)+(j-1)*m+1:m*(i-1)*
137              (M-1)+(j-1)*m+m,1)=u((i-1)*M*m+j*m+1:(i-1)*M*m+j*m+m);
138              end
139          end
140          j=1;
141          for i=temp*M:N-1
142              conAttck(m*(temp-1)*(M-1)+(M-1-1)*m+m+(j-1)*m+1:m*(temp-1)*
143              (M-1)+(M-1-1)*m+m+j*m)=u(temp*M*m+1:temp*M*m+m);
144              ucon(m*(temp-1)*(M-1)+(M-1-1)*m+m+(j-1)*m+1:m*(temp-1)*
145              (M-1)+(M-1-1)*m+m+j*m)=u(i*m+1:i*m+m);
146              j=j+1;
147          end
148
149              con=con+(ucon==conAttck);
150          end
151          con=con+(u(:)≤uh);
152          con=con+(u(:)≥ul);
153          con=con+(u≤uit+.1);
154          con=con+(u≥uit-.1);
155      end
156
157      end
```

Bibliography

[Ahmad *et al.* 2018] S. Ahmad, A. Badwelan, A. M. Ghaleb, A. Qamhan, M. Sharaf, M. Alatefi and A. Moohialdin. *Analyzing Critical Failures in a Production Process: Is Industrial IoT the Solution?* Wireless Communications and Mobile Computing, 2018. (Cited on page 13.)

[Alcácer & Cruz-Machado 2019] V. Alcácer and V. Cruz-Machado. *Scanning the Industry 4.0: A Literature Review on Technologies for Manufacturing Systems.* Engineering Science and Technology, an International Journal, vol. 22, no. 3, pages 899 – 919, 2019. (Cited on page 14.)

[Alladi *et al.* 2020] T. Alladi, V. Chamola and S. Zeadally. *Industrial Control Systems: Cyberattack trends and countermeasures.* Computer Communications, vol. 155, pages 1 – 8, 2020. (Cited on page 16.)

[Amin *et al.* 2009] Saurabh Amin, Alvaro A. Cárdenas and S. Shankar Sastry. *Safe and Secure Networked Control Systems under Denial-of-Service Attacks.* In Rupak Majumdar and Paulo Tabuada, editors, Hybrid Systems: Computation and Control, pages 31–45, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. (Cited on page 15.)

[Aps 2020] Mosek Aps. *Mosek optimization toolbox for Matlab - Release 9.2.29.* Technical report, 2020. (Cited on page 60.)

[Archela *et al.* 2018] A. Archela, D. G. Toginho and L. F. d. Melo. *Torque Control of a DC Motor With a State Space Estimator and Kalman Filter for Vehicle Traction.* In 2018 13th IEEE International Conference on Industry Applications (INDUSCON), pages 763–769, 2018. (Cited on pages 96 and 109.)

[Atallah *et al.* 2007] K. Atallah, S. D. Calverley and D. Howe. *A Brushless Permanent Magnet Motor With Integrated Torque-Limiter.* IEEE Transactions on Magnetics, vol. 43, no. 6, pages 2498–2500, 2007. (Cited on page 42.)

[Bai *et al.* 2017a] Cheng-Zong Bai, Vijay Gupta and Fabio Pasqualetti. *On Kalman Filtering with Compromised Sensors: Attack Stealthiness and Performance Bounds.* IEEE Trans. Autom. Control, vol. 62, no. 12, pages 6641–6648, December 2017. (Cited on page 38.)

[Bai *et al.* 2017b] Cheng-Zong Bai, Fabio Pasqualetti and Vijay Gupta. *Data-injection attacks in stochastic control systems: Detectability and performance tradeoffs.* Automatica, vol. 82, pages 251–260, August 2017. (Cited on page 39.)

[Bartnikas & Morin 2004] R. Bartnikas and R. Morin. *Multi-stress aging of stator bars with electrical, thermal, and mechanical stresses as simultaneous acceleration factors.* IEEE Transactions on Energy Conversion, vol. 19, no. 4, pages 702–714, 2004. (Cited on page 42.)

[Bazrafshan *et al.* 2013] Z. Bazrafshan, H. Hashemi, S. M. H. Fard and A. Hamzeh. *A survey on heuristic malware detection techniques.* In The 5th Conference on Information and Knowledge Technology, pages 113–120, 2013. (Cited on page 32.)

[Beaudet *et al.* 2020] A. Beaudet, F. Sicard, C. Escudero and E. Zamaï. *Process-Aware Model-based Intrusion Detection System on Filtering Approach: Further Investigations.* In 2020 IEEE International Conference on Industrial Technology (ICIT), pages 310–315, 2020. (Cited on page 33.)

[Beaudet *et al.* 2021] A. Beaudet, C. Escudero and E. Zamaï. *Malicious Anomaly Detection Approaches Robustness in Manufacturing ICSs.* In Proc. INCOM, **(Submitted)**, pages 1–8, Budapest, June 2021. (Not cited.)

[Ben-Talha *et al.* 2017] H. Ben-Talha, P. Massioni and G. Scorletti. *Robust simulation of continuous-time systems with rational dynamics.* International Journal of Robust and Nonlinear Control, vol. 27, no. 16, pages 3097–3108, 2017. (Cited on page 81.)

[Blanchini & Miani 2015] Franco Blanchini and Stefano Miani. Springer International Publishing, Cham, 2015. (Cited on pages 45, 55, 57 and 71.)

[Blanchini 1999] F. Blanchini. *Set invariance in control.* Automatica, vol. 35, no. 11, pages 1747 – 1767, 1999. (Cited on pages 45 and 55.)

[Bolton 2015] W. Bolton. Programmable logic controllers (sixth edition). Newnes, Boston, sixth edition édition, 2015. (Cited on pages 20 and 22.)

[Bouaziz 2012] E. Bouaziz M.-F.and Zamaï. *Equipment Health Factor prediction for complex semiconductor manufacturing facility.* IFAC Proceedings Volumes, vol. 45, no. 6, pages 1005 – 1010, 2012. 14th IFAC Symposium on Information Control Problems in Manufacturing. (Cited on page 12.)

[Boyd & Grant 2020] Stephen Boyd and Michael C. Grant. *The CVX users' guide - Release 2.2.* Technical report, 2020. (Cited on page 60.)

[Boyd & Vandenberghe 2004] Stephen Boyd and Lieven Vandenberghe. Convex optimization. Cambridge University Press, USA, 2004. (Cited on pages 56, 58 and 59.)

[Boyd *et al.* 1994] S.P. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan. Linear matrix inequalities in system and control theory, volume 15. SIAM, 1994. (Cited on pages 52, 53 and 55.)

[Bradbury 2012] Danny Bradbury. *SCADA: a critical vulnerability.* Computer Fraud & Security, vol. 2012, no. 4, pages 11 – 14, 2012. (Cited on page 22.)

[Cagalaban *et al.* 2010] Giovanni Cagalaban, Taihoon Kim and Seoksoo Kim. *Improving SCADA Control Systems Security with Software Vulnerability Analysis.* In Proceedings of the 12th WSEAS International Conference on Automatic Control, Modelling & Simulation, page 409–414, Stevens Point, Wisconsin, USA, 2010. World Scientific and Engineering Academy and Society (WSEAS). (Cited on page 23.)

[Carcano *et al.* 2009] Andrea Carcano, Igor Nai Fovino, Marcelo Masera and Alberto Trombetta. *State-based network intrusion detection systems for SCADA protocols: a proof of concept.* In International Workshop on Critical Information Infrastructures Security, pages 138–150. Springer, 2009. (Cited on page 32.)

[Carfagno *et al.* 1980] S.P. Carfagno, R.J. Gibson, Franklin Research Center and Electric Power Research Institute. A review of equipment aging theory and technology. EPRI. Electric Power Research Institute, 1980. (Cited on page 41.)

[Caselli *et al.* 2015] Marco Caselli, Emmanuele Zambon and Frank Kargl. *Sequence-aware Intrusion Detection in Industrial Control Systems.* In Proceedings of the 1st ACM Workshop on Cyber-Physical System Security, pages 13–24, New York, NY, USA, 2015. ACM. (Cited on page 33.)

[Cetinkaya *et al.* 2019] Ahmet Cetinkaya, Hideaki Ishii and Tomohisa Hayakawa. *An Overview on Denial-of-Service Attacks in Control Systems: Attack Models and Security Analyses.* Entropy, vol. 21, page 210, 02 2019. (Cited on page 15.)

[Chang *et al.* 2018] Young Hwan Chang, Qie Hu and Claire J. Tomlin. *Secure estimation based Kalman Filter for cyber–physical systems against sensor attacks.* Automatica, vol. 95, pages 399–412, September 2018. (Cited on page 39.)

[Chen *et al.* 2004] Chieh-Li Chen, Kuo-Ming Chang and Chih-Ming Chang. *Modeling and control of a web-fed machine.* Appl. Math. Model., vol. 28, no. 10, pages 863 – 876, Oct. 2004. (Cited on pages vii, 96, 97, 98 and 99.)

[Cheng *et al.* 2019] P. Cheng, Z. Yang, J. Chen, Y. Qi and L. Shi. *An event-based stealthy attack on remote state estimation.* IEEE Trans. Autom. Control, pages 1–1, 2019. (Cited on page 38.)

[Cherubini *et al.* 2016] Andrea Cherubini, Robin Passama, André Crosnier, Antoine Lasnier and Philippe Fraisse. *Collaborative manufacturing with physical human–robot interaction.* Robotics and Computer-Integrated Manufacturing, vol. 40, pages 1 – 13, 2016. (Cited on page 8.)

[Chesi 2010] G. Chesi. *LMI techniques for optimization over polynomials in control: a survey.* IEEE Transactions on Automatic Control, vol. 55, no. 11, pages 2500–2510, 2010. (Cited on pages 54 and 55.)

[Cómbita *et al.* 2019] L. F. Cómbita, Á. A. Cárdenas and N. Quijano. *Mitigating Sensor Attacks Against Industrial Control Systems.* IEEE Access, vol. 7, pages 92444–92455, 2019. (Cited on page 43.)

[Combacau *et al.* 2000] M. Combacau, P. Berruet, E. Zamai, P. Charbonnaud and A. Khatab. *Supervision and Monitoring of Production Systems.* In Proc. IFAC MCPL (IFAC Proceedings Volumes), volume 33, pages 849 – 854, 2000. (Cited on page 24.)

[Cox *et al.* 1995] D.F. Cox, K.L. McElhaney and R.H. Staunton. Aging and service wear of air-operated valves used in safety-related systems at nuclear power plants. Number 6016 de NUREG Series. International Atomic Energy Agency, United States, 1995. (Cited on page 41.)

[Dai & Vyatkin 2012] W. Dai and V. Vyatkin. *Redesign Distributed PLC Control Systems Using IEC 61499 Function Blocks.* IEEE Trans. Autom. Sci. Eng, vol. 9, no. 2, pages 390–401, 2012. (Cited on page 19.)

[Daneels & Salter 1999] A. Daneels and W. Salter. *What is SCADA?* In D. Bulfone and A. Daneels, editors, 7th International Conference on Accelerator and Large Experimental Physics Control Systems (ICALEPCS 99), volume 991004, pages 339–343, 1999. (Cited on page 23.)

[de Sá *et al.* 2017] A. O. de Sá, L. F. R. d. C. Carmo and R. C. S. Machado. *Covert Attacks in Cyber-Physical Control Systems.* IEEE Trans. Ind. Informat., vol. 13, no. 4, pages 1641–1651, Aug. 2017. (Cited on pages 15, 43 and 46.)

[Dhillon 1999] B.S. Dhillon. Design reliability: Fundamentals and applications. CRC Press, 1999. (Cited on page 12.)

[Doggett 2004] A. M. Doggett. *A Statistical Comparison of Three Root Cause Analysis Tools.* Journal of Industrial Technology, vol. 20, no. 2, pages 2 – 9, 2004. (Cited on page 12.)

[Dorf & Bishop 2000] Richard C. Dorf and Robert H. Bishop. Modern control systems. Prentice-Hall, Inc., USA, 9th édition, 2000. (Cited on page 36.)

[Dorr *et al.* 1997] R. Dorr, F. Kratz, J. Ragot, F. Loisy and J. . Germain. *Detection, isolation, and identification of sensor faults in nuclear power plants.* IEEE Transactions on Control Systems Technology, vol. 5, no. 1, pages 42–60, 1997. (Cited on page 24.)

[Duong *et al.* 2013] Quoc-Bao Duong, Eric Zamai and Khoi-Quoc Tran-Dinh. *Confidence estimation of feedback information for logicdiagnosis.* Engineering Applications of Artificial Intelligence, vol. 26, no. 3, pages 1149 – 1161, 2013. (Cited on page 11.)

[Erez & Wool 2015] Noam Erez and Avishai Wool. *Control variable classification, modeling and anomaly detection in Modbus/TCP SCADA systems.* International Journal of Critical Infrastructure Protection, vol. 10, pages 59–70, September 2015. (Cited on page 32.)

[Erickson & Hedrick 1999] K.T. Erickson and J.L. Hedrick. Plantwide process control. Wiley Series in Chemical Engineering. Wiley, 1999. (Cited on page 12.)

[Escudero & Zamaï 2019] C. Escudero and E. Zamaï. *Prevention of Aging Attacks: Malicious Nature of the Control Signal.* In 2019 International Automatic Control Conference (CACS), pages 1–6, Nov 2019. (Cited on page 70.)

[Escudero *et al.* 2018] C. Escudero, F. Sicard and E. Zamai. *Process-Aware Model based IDSs for Industrial Control Systems Cybersecurity: Approaches, Limits and Further Research.* In 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), volume 1, pages 605–612, Sep. 2018. (Cited on page 15.)

[Escudero *et al.* 2020a] C. Escudero, P. Massioni, G. Scorletti and E. Zamaï. *Security of Control Systems: Prevention of Aging Attacks by means of Convex Robust Simulation Forecasts.* Proc. IFAC WC (IFAC-PapersOnLine), pages 1 – 1, Jul. 2020. (Cited on page 80.)

[Escudero *et al.* 2020b] C. Escudero, P. Massioni, E. Zamaï and B. Raison. *Control System Theory Methods to Analyze, Prevent and Design Stealthy Aging Attacks.* IEEE Transaction on Industrial Informatics, **(Submitted)**, pages 1 – 1, Sept. 2020. (Not cited.)

[Escudero *et al.* 2020c] C. Escudero, F. Sicard, A. Beaudet and E. Zamaï. *Security of the Process Integrity under Attacks Propagation in Industrial Control Systems: a Classification of Control-based Methods.* Reliab. Eng. Syst. Saf., **(Submitted)**, pages 1 – 1, Sept. 2020. (Not cited.)

[Fawzi *et al.* 2014] Hamza Fawzi, Paulo Tabuada and Suhas Diggavi. *Secure Estimation and Control for Cyber-Physical Systems Under Adversarial Attacks.* IEEE Trans. Autom. Control, vol. 59, no. 6, pages 1454–1467, June 2014. (Cited on page 39.)

[Fovino *et al.* 2010] I. N. Fovino, A. Carcano, T. D. L. Murel, A. Trombetta and M. Masera. *Modbus/DNP3 State-Based Intrusion Detection System.* In 2010 24th IEEE International Conference on Advanced Information Networking and Applications, pages 729–736, April 2010. (Cited on page 32.)

[Fovino 2014] Igor Nai Fovino. Scada system cyber security, pages 451–471. Springer New York, New York, NY, 2014. (Cited on pages 16 and 23.)

[Fuc 2008] *Chapter 1 - Introduction to Power Quality.* In E. F. Fuchs and M. A.S. Masoum, editors, Power Quality in Power Systems and Electrical Machines, pages 1 – 54. Academic Press, Burlington, 2008. (Cited on page 42.)

[Galloway & Hancke 2013] B. Galloway and G. P. Hancke. *Introduction to Industrial Control Networks.* IEEE Communications Surveys & Tutorials, vol. 15, no. 2, pages 860–880, 2013. (Cited on page 13.)

[Garcia *et al.* 2016] L. Garcia, S. Zonouz, Dong Wei and L. P. de Aguiar. *Detecting PLC control corruption via on-device runtime verification.* In 2016 Resilience Week (RWS), pages 67–72, August 2016. (Cited on page 33.)

[Gayek 1991] J. E. Gayek. *A survey of techniques for approximating reachable and controllable sets.* In Proc. CDC, pages 1724–1729, Dec. 1991. (Cited on pages 45 and 70.)

[Gisel & Olejnik 2018] L. Gisel and L. Olejnik. *The potential human cost of cyber operations.* Technical report, International Committee of the Red Cross Expert Meeting, Nov. 2018. (Cited on page 14.)

[Globalspec] Globalspec. *Drive rollers and idler rollers information.* https://www.globalspec.com/learnmore/material_handling_packaging_equipment/material_handling_equipment/conveyor_rollers. [Online; accessed 19-November-2020]. (Cited on page 96.)

[Goldenberg & Wool 2013] Niv Goldenberg and Avishai Wool. *Accurate modeling of Modbus/TCP for intrusion detection in SCADA systems.* International Journal of Critical Infrastructure Protection, vol. 6, no. 2, pages 63–75, June 2013. (Cited on page 32.)

[Gonzalez *et al.* 2019] D. Gonzalez, F. Alhenaki and M. Mirakhorli. *Architectural Security Weaknesses in Industrial Control Systems (ICS) an Empirical Study Based on Disclosed Software Vulnerabilities.* In 2019 IEEE International Conference on Software Architecture (ICSA), pages 31–40, 2019. (Cited on pages 22 and 23.)

[Groover 2008] M.P. Groover. Automation, production systems, and computer-integrated manufacturing. Prentice Hall, 2008. (Cited on page 8.)

[Gungor & Hancke 2009] V. C. Gungor and G. P. Hancke. *Industrial Wireless Sensor Networks: Challenges, Design Principles, and Technical Approaches.* IEEE Transactions on Industrial Electronics, vol. 56, no. 10, pages 4258–4265, 2009. (Cited on page 14.)

[Guo *et al.* 2018] Ziyang Guo, Dawei Shi, Karl Henrik Johansson and Ling Shi. *Worst-case stealthy innovation-based linear attack on remote state estimation.* Automatica, vol. 89, pages 117–124, March 2018. (Cited on page 38.)

[Guo *et al.* 2019] Ziyang Guo, Dawei Shi, Karl H. Johansson and Ling Shi. *Worst-case Innovation-based Integrity Attacks with Side Information on Remote State Estimation.* IEEE Control Netw. Syst., vol. 6, no. 1, pages 48–59, March 2019. (Cited on page 38.)

[Haapala *et al.* 2010] Antti Haapala, Henrikki Liimatainen, Mika Körkkö, Jaakko Ekman, Mirja Sinikka Salkinoja-Salonen and Jouko Niinimäki. *Web defects in newsprint production –a mill case study.* Appita Journal, vol. 63, no. 5, pages 358–364, 2010. (Cited on page 13.)

[Hadizadeh Kafash *et al.* 2018] S. Hadizadeh Kafash, N. Hashemi, C. Murguia and J. Ruths. *Constraining Attackers and Enabling Operators via Actuation Limits.* In 2018 IEEE Conference on Decision and Control (CDC), pages 4535–4540, Dec 2018. (Cited on page 40.)

[Hadziosmanovic *et al.* 2013] Dina Hadziosmanovic, Robin Sommer, Emmanuele Zambon and Pieter Hartel. *Through the eye of the PLC: towards semantic security*

monitoring for industrial control systems. International Computer Science Institute, Berkeley, 2013. (Cited on page 32.)

[Hahn & Govindarasu 2013] Adam Hahn and Manimaran Govindarasu. *Model-based Intrustion Detection for the Smart Grid (MINDS).* In Proceedings of the Eighth Annual Cyber Security and Information Intelligence Research Workshop, pages 27:1–27:4, New York, NY, USA, 2013. ACM. (Cited on page 32.)

[Han & Palazzolo 2013] X. Han and A. B. Palazzolo. *VFD Machinery Vibration Fatigue Life and Multilevel Inverter Effect.* IEEE Transactions on Industry Applications, vol. 49, no. 6, pages 2562–2575, 2013. (Cited on page 42.)

[Hassibi & Boyd 1998] A. Hassibi and S. Boyd. *Quadratic stabilization and control of piecewise-linear systems.* In Proceedings of the 1998 American Control Conference (ACC), volume 6, pages 3659–3664 vol.6, 1998. (Cited on page 37.)

[Hemsley & Fisher 2018a] K. E. Hemsley and R. E. Fisher. *History of Industrial Control System Cyber Incidents.* Technical report, Idaho National Laboratory, Dec. 2018. (Cited on pages 14 and 15.)

[Hemsley & Fisher 2018b] Kevin Hemsley and Ronald Fisher. *A History of Cyber Incidents and Threats Involving Industrial Control Systems.* In Jason Staggs and Sujeet Shenoi, editors, Critical Infrastructure Protection XII, pages 215–242, Cham, 2018. Springer International Publishing. (Cited on page 16.)

[Holopainen *et al.* 2013] T. P. Holopainen, P. Jorg, J. Nïiranen and D. Andrea. *Electric Motors and Drives in Torsional Vibration Analysis and Design.* In Proceedings of the 42nd Turbomachinery Symposium, 2013. (Cited on page 42.)

[Hoo *et al.* 2003] K. A. Hoo, M. J. Piovoso, P. D. Schnelle and D. A. Rowan. *Process and controller performance monitoring: overview with industrial applications.* International Journal of Adaptive Control and Signal Processing, vol. 17, no. 7-9, pages 635–662, 2003. (Cited on page 21.)

[Hristopulos & Uesaka 2002] D.T. Hristopulos and T. Uesaka. *A model of machine-direction tension variations in paper webs with runnability applications.* Journal of Pulp and Paper Science, vol. 28, no. 12, pages 389 – 394, Dec. 2002. (Cited on page 12.)

[Hu *et al.* 2018a] Liang Hu, Zidong Wang, Qing-Long Han and Xiaohui Liu. *State estimation under false data injection attacks: Security analysis and system protection.* Automatica, vol. 87, pages 176–183, January 2018. (Cited on page 39.)

[Hu *et al.* 2018b] Yan Hu, Hong Li, Tom H. Luan, An Yang, Limin Sun, Zhiliang Wang and Rui Wang. *Detecting stealthy attacks on industrial control systems using a permutation entropy-based method.* Future Generation Computer Systems, August 2018. (Cited on page 33.)

[Hubac & Zamaï 2013] S. Hubac and E. Zamaï. *Politiques de maintenance équipement en flux de production stressant.* Techniques de l'ingénieur Méthodes de production, vol. base documentaire : TIB521DUO., 2013. (Cited on page 13.)

[IAE 1989] Safety aspects of nuclear power plant ageing (state of the art report). Number 540 de TECDOC Series. International Atomic Energy Agency, Vienna, 1989. (Cited on page 42.)

[ISA] ISA. *ISA 95 Standard.* https://isa-95.com/. [Online; accessed 19-November-2020]. (Cited on page 10.)

[Isermann 1997] R. Isermann. *Supervision, fault-detection and fault-diagnosis methods — An introduction.* Control Engineering Practice, vol. 5, no. 5, pages 639 – 652, 1997. (Cited on page 11.)

[Isermann 2006] R. Isermann. Fault-diagnosis systems: An introduction from fault detection to fault tolerance. Springer-Verlag Berlin Heidelberg, 2006. (Cited on page 12.)

[Ishikawa 1990] K. Ishikawa. Introduction to quality control. Taylor & Francis, 1990. (Cited on page 12.)

[ISO13849-1 2015] ISO13849-1. 13849-1:2015: Safety of machinery, safety related parts of control systems, part 1: General principles for design. 3rd edition. International Organization for Standardization, 2015. (Cited on page 11.)

[Jones & McLean 1986] Albert T. Jones and Charles R. McLean. *A proposed hierarchical control model for automated manufacturing systems.* J. Manuf. Syst., vol. 5, no. 1, pages 15 – 25, 1986. (Cited on page 20.)

[Jones & Saleh 1989] A. Jones and A. Saleh. *A multi-layer/multi-level control architecture for computer integrated manufacturing systems.* In 15th Annual Conference of IEEE Industrial Electronics Society, pages 519–525 vol.3, 1989. (Cited on page 10.)

[Juliani Correa de Godoy & Garcia 2017] Rodrigo Juliani Correa de Godoy and Claudio Garcia. *Plantwide Control: A Review of Design Techniques, Benchmarks, and Challenges.* Industrial & Engineering Chemistry Research, vol. 56, no. 28, pages 7877–7887, 2017. (Cited on page 23.)

[Kafash *et al.* 2018] S. H. Kafash, J. Giraldo, C. Murguia, A. A. Cardenas and J. Ruths. *Constraining Attacker Capabilities Through Actuator Saturation.* In Proc. ACC, pages 986–991, Milwaukee, WI, Jun. 2018. (Cited on pages 40 and 45.)

[Kalech 2019] Meir Kalech. *Cyber-attack detection in SCADA systems using temporal pattern recognition techniques.* Computers & Security, vol. 84, pages 225–238, July 2019. (Cited on page 32.)

[Kalman & Bertram 1960] R.E. Kalman and J.E. Bertram. *Control system analysis and design via the second method of Lyapunov: Continuous-time systems.* Journal of Basic Engineering, vol. 82, no. 2, pages 371–393, 1960. (Cited on page 61.)

[Khlebnikov *et al.* 2011] M. V. Khlebnikov, B. T. Polyak and V. M. Kuntsevich. *Optimization of linear systems subject to bounded exogenous disturbances: The invariant ellipsoid technique.* Autom Remote Control, vol. 72, no. 11, pages 2227 – 2275, 2011. (Cited on page 45.)

[Kidam & Hurme 2013] K. Kidam and M. Hurme. *Analysis of equipment failures as contributors to chemical process accidents.* Process Safety and Environmental Protection, vol. 91, no. 1, pages 61 – 78, 2013. (Cited on page 13.)

[Kim 2012] HyungJun Kim. *Security and Vulnerability of SCADA Systems over IP-Based Wireless Sensor Networks.* International Journal of Distributed Sensor Networks, vol. 8, no. 11, page 268478, 2012. (Cited on page 23.)

[Kleinmann & Wool 2017] Amit Kleinmann and Avishai Wool. *Automatic Construction of Statechart-Based Anomaly Detection Models for Multi-Threaded Industrial Control Systems.* ACM Trans. Intell. Syst. Technol., vol. 8, no. 4, pages 55:1–55:21, February 2017. (Cited on page 32.)

[Kobbacy & Murthy 2008] Khairy A. H Kobbacy and D. N. P Murthy. Complex system maintenance handbook. Springer London, 1st ed. édition, 2008. (Cited on page 12.)

[Koucham *et al.* 2018] Oualid Koucham, Stéphane Mocanu, Guillaume Hiet, Jean-Marc Thiriet and Frédéric Majorczyk. *Efficient Mining of Temporal Safety Properties for Intrusion Detection in Industrial Control Systems.* In Proc. IFAC SAFEPROCESS, volume 51, pages 1043 – 1050, Warsaw, Aug. 2018. (Cited on page 33.)

[Koutsandria *et al.* 2014] G. Koutsandria, V. Muthukumar, M. Parvania, S. Peisert, C. McParland and A. Scaglione. *A hybrid network IDS for protective digital relays in the power transmission grid.* In 2014 IEEE International Conference on Smart Grid Communications (SmartGridComm), pages 908–913, November 2014. (Cited on page 33.)

[Kramer *et al.* 1984] J. Kramer, J. Magee and M. Sloman. *A software architecture for distributed computer control systems.* Automatica, vol. 20, no. 1, pages 93 – 102, 1984. (Cited on page 20.)

[Kurzhanskiy & Varaiya 2006] A. A. Kurzhanskiy and P. Varaiya. Ellipsoidal toolbox. EECS Department, University of California, Berkeley, Tech. Rep., 2006. (Cited on page 58.)

[Lalanne 2013] C. Lalanne. Mechanical vibration and shock analysis, fatigue damage. Mechanical Vibration and Shock Analysis. Wiley, 2013. (Cited on page 42.)

[Langner 2011] R. Langner. *Stuxnet: Dissecting a Cyberwarfare Weapon.* IEEE Secur Priv, vol. 9, no. 3, pages 49–51, 2011. (Cited on page 28.)

[Li & Yang 2019] Yi-Gang Li and Guang-Hong Yang. *Optimal Stealthy False Data Injection Attacks in Cyber-Physical Systems.* Inform. Sciences, vol. 481, pages 474–490, May 2019. (Cited on page 38.)

[Li *et al.* 2019] W. Li, L. Xie and Z. Wang. *Two-Loop Covert Attacks Against Constant Value Control of Industrial Control Systems.* IEEE Trans. Ind. Informat., vol. 15, no. 2, pages 663–676, Feb 2019. (Cited on page 40.)

[Lin *et al.* 2013] Hui Lin, Adam Slagell, Catello Di Martino, Zbigniew Kalbarczyk and Ravishankar K. Iyer. *Adapting Bro into SCADA: Building a Specification-based Intrusion Detection System for the DNP3 Protocol.* In Proceedings of the Eighth Annual Cyber Security and Information Intelligence Research Workshop, CSIIRW '13, pages 5:1–5:4, New York, NY, USA, 2013. ACM. (Cited on page 32.)

[Linda *et al.* 2009] O. Linda, T. Vollmer and M. Manic. *Neural Network based Intrusion Detection System for critical infrastructures.* In 2009 International Joint Conference on Neural Networks, pages 1827–1834, June 2009. (Cited on page 32.)

[Liu *et al.* 2014] C. Liu, D. Jiang and J. Chen. *Coupled Torsional Vibration and Fatigue Damage of Turbine Generator Due to Grid Disturbance.* Journal of Engineering for Gas Turbines and Power, vol. 136, Jun. 2014. (Cited on page 42.)

[Liu *et al.* 2019] Y. Liu, M. Kashef, K. B. Lee, L. Benmohamed and R. Candell. *Wireless Network Design for Emerging IIoT Applications: Reference Framework and Use Cases.* Proceedings of the IEEE, vol. 107, no. 6, pages 1166–1192, 2019. (Cited on page 14.)

[Lofberg 2004] J. Lofberg. *YALMIP : a toolbox for modeling and optimization in MATLAB.* In 2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508), pages 284–289, 2004. (Cited on page 60.)

[Löfberg 2009] J. Löfberg. *Pre- and post-processing sum-of-squares programs in practice.* IEEE Transactions on Automatic Control, vol. 54, no. 5, pages 1007–1011, 2009. (Cited on page 54.)

[Long *et al.* 2005] M. Long, Chwan-Hwa Wu and J. Y. Hung. *Denial of service attacks on network-based control systems: impact and mitigation.* IEEE Transactions on Industrial Informatics, vol. 1, no. 2, pages 85–96, 2005. (Cited on page 15.)

[Lucchetti 2006] Roberto Lucchetti. Convex sets and convex functions: the fundamentals, pages 1–19. Springer New York, New York, NY, 2006. (Cited on pages 55 and 56.)

[Luiijf 2016] E. Luiijf. Threats in industrial control systems, pages 69–93. Springer International Publishing, 2016. (Cited on page 16.)

[Lyashevskiy & Meyer 1995] S. Lyashevskiy and A. U. Meyer. *Control system analysis and design upon the Lyapunov method.* In Proceedings of 1995 American Control Conference - ACC'95, volume 5, pages 3219–3223 vol.5, 1995. (Cited on page 61.)

[M-SYSTEM 1990] M-SYSTEM. *SC100/200 Series Function Block Application Manual EM-6460-C Rev.1.* Technical report, 1990. (Cited on page 19.)

[Madonna *et al.* 2020] V. Madonna, P. Giangrande, L. Lusuardi, A. Cavallini, C. Gerada and M. Galea. *Thermal Overload and Insulation Aging of Short Duty Cycle, Aerospace Motors.* IEEE Transactions on Industrial Electronics, vol. 67, no. 4, pages 2618–2629, 2020. (Cited on page 42.)

[Mano 1992] M. Morris Mano. Computer system architecture (3rd edition), pages 1–4. Prentice Hall, 3 édition, 1992. (Cited on page 20.)

[Marelli *et al.* 2018] Damián Marelli, Tianju Sui and Minyue Fu. *Statistical Approach to Detection of Attacks for Stochastic Cyber-Physical Systems.* IFAC-PapersOnLine, vol. 51, no. 25, pages 178–183, January 2018. (Cited on page 33.)

[McGrew & Vaughn 2009] Robert Wesley McGrew and Rayford B. Vaughn. *Discovering vulnerabilities in control system human–machine interface software.* Journal of Systems and Software, vol. 82, no. 4, pages 583 – 589, 2009. (Cited on page 23.)

[Mcgrew 2013] Robert Wesley Mcgrew. *Vulnerability Analysis Case Studies of Control Systems Human Machine Interfaces.* PhD thesis, USA, 2013. (Cited on page 23.)

[McLaughlin 2013] Stephen McLaughlin. *CPS: Stateful Policy Enforcement for Control System Device Usage.* In Proceedings of the 29th Annual Computer Security Applications Conference, ACSAC '13, pages 109–118, New York, NY, USA, 2013. ACM. (Cited on page 32.)

[McNeill 1983] H McNeill. *Environmental qualification of essential safety-related equipment for nuclear power plants.* Nucl. Saf., vol. 24:5, pages 658–670, Oct. 1983. (Cited on page 42.)

[Merouane *et al.* 2020] E. M. Merouane, C. Escudero, F. Sicard and E. Zamai. *Aging Attacks against Electro-Mechanical Actuators from Control Signal Manipulation.* In 2020 IEEE International Conference on Industrial Technology (ICIT), pages 133–138, 2020. (Cited on page 42.)

[Michalos *et al.* 2010] G. Michalos, S. Makris, N. Papakostas, D. Mourtzis and G. Chryssolouris. *Automotive assembly technologies review: challenges and outlook for a flexible and adaptive approach.* CIRP Journal of Manufacturing Science and Technology, vol. 2, no. 2, pages 81 – 91, 2010. (Cited on pages 8 and 12.)

[Milinković & Lazić 2012] S. A. Milinković and L. R. Lazić. *Industrial PLC security issues.* In 2012 20th Telecommunications Forum (TELFOR), pages 1536–1539, 2012. (Cited on page 22.)

[Miloševič *et al.* 2017] Jezdimir Miloševič, Takashi Tanaka, Henrik Sandberg and Karl Henrik Johansson. *Analysis and Mitigation of Bias Injection Attacks Against a Kalman Filter.* In Proc. IFAC World Congress, pages 8393–8398, July 2017. (Cited on page 39.)

[Milošević *et al.* 2019] J. Milošević, H. Sandberg and K. H. Johansson. *Estimating the Impact of Cyber-Attack Strategies for Stochastic Networked Control Systems.* IEEE Control Netw. Syst., pages 1–1, 2019. (Cited on page 39.)

[Mitchell *et al.* 2005] I. M. Mitchell, A. M. Bayen and C. J. Tomlin. *A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games.* IEEE Transactions on Automatic Control, vol. 50, no. 7, pages 947–957, 2005. (Cited on page 45.)

[Mo & Sinopoli 2016a] Y. Mo and B. Sinopoli. *On the Performance Degradation of Cyber-Physical Systems Under Stealthy Integrity Attacks.* IEEE Transactions on Automatic Control, vol. 61, no. 9, pages 2618–2624, September 2016. (Cited on page 33.)

[Mo & Sinopoli 2016b] Yilin Mo and Bruno Sinopoli. *On the Performance Degradation of Cyber-Physical Systems Under Stealthy Integrity Attacks.* IEEE Trans. Autom. Control, vol. 61, no. 9, pages 2618–2624, September 2016. (Cited on page 39.)

[Morris *et al.* 2013] Thomas H. Morris, Rayford B. Vaughn and Elena Sitnikova. *Advances in the Protection of Critical Infrastructure by Improvement in Industrial Control System Security.* In Proceedings of the Eleventh Australasian Information Security Conference - Volume 138, page 67–73, 2013. (Cited on page 23.)

[Muñoz *et al.* 2011] E. Muñoz, E. Capón-García, M. Moreno-Benito, A. Espuña and L. Puigjaner. *Scheduling and control decision-making under an integrated information environment.* Computers & Chemical Engineering, vol. 35, no. 5, pages 774 – 786, 2011. (Cited on pages 12 and 13.)

[Murguia & Ruths 2019] C. Murguia and J. Ruths. *On model-based detectors for linear time-invariant stochastic systems under sensor attacks.* IET Control Theory Appl., vol. 13, no. 8, pages 1051–1061, May 2019. (Cited on page 39.)

[Murguia *et al.* 2017] Carlos Murguia, Nathan van de Wouw and Justin Ruths. *Reachable Sets of Hidden CPS Sensor Attacks: Analysis and Synthesis Tools.* IFAC-PapersOnLine, vol. 50, no. 1, pages 2088–2094, July 2017. (Cited on page 33.)

[Murguia *et al.* 2020a] Carlos Murguia, Iman Shames, Justin Ruths and Dragan Nešić. *Security metrics and synthesis of secure control systems.* Automatica, vol. 115, page 108757, May 2020. (Cited on page 33.)

[Murguia *et al.* 2020b] Carlos Murguia, Iman Shames, Justin Ruths and Dragan Nešić. *Security metrics and synthesis of secure control systems.* Automatica, vol. 115, page 108757, May 2020. (Cited on pages 38 and 45.)

[Nature] Nature. *Dynamical systems.* https://www.nature.com/subjects/dynamical-systems. [Online; accessed September 15, 2020]. (Cited on page 36.)

[Nguyen *et al.* 2016] D. T. Nguyen, Q. B. Duong, E. Zamaï and Muhammad Kashif Shahzad. *Fault diagnosis for the complex manufacturing system.* Proc Inst Mech Eng O J Risk Reliab, vol. 230, no. 2, pages 178–194, 2016. (Cited on page 12.)

[Nguyen 2015] Dang-Trinh Nguyen. *Diagnostic en ligne des systèmes à événements discrets complexes : approche mixte logique/probabiliste*. PhD thesis, 2015. Thèse de doctorat dirigée par Zamaï, Eric Automatique et productique Université Grenoble Alpes (ComUE) 2015. (Cited on page 32.)

[Nozaic & Freese 2009] DJ Nozaic and SD Freese. *Process Design Manual For Small Wastewater Works*. Technical report, Water Research Commission, Apr. 2009. (Cited on page 8.)

[OECD 2017] OECD. *Enhancing the Role of Insurance in Cyber Risk Management*. Technical report, Dec. 2017. (Cited on page 14.)

[Palacios-Gomez *et al.* 1982] F. Palacios-Gomez, L. Lasdon and M. Engquist. *Nonlinear Optimization by Successive Linear Programming*. Management Science, vol. 28, no. 10, pages 1106–1120, 1982. (Cited on page 91.)

[Parrilo 2003] P.A. Parrilo. *Semidefinite programming relaxations for semialgebraic problems*. Mathematical programming, vol. 96, no. 2, pages 293–320, 2003. (Cited on page 54.)

[Pavesi *et al.* 2019] Jaime Pavesi, Thamara Villegas, Alexey Perepechko, Eleazar Aguirre and Lorena Galeazzi. *Validation of ICS Vulnerability Related to TCP/IP Protocol Implementation in Allen-Bradley Compact Logix PLC Controller*. In Miguel Felix Mata-Rivera, Roberto Zagal-Flores and Cristian Barría-Huidobro, editors, Telematics and Computing, pages 355–364, Cham, 2019. Springer International Publishing. (Cited on page 22.)

[Pecsvaradi & Narendra 1971] Thomas Pecsvaradi and Kumpati S. Narendra. *Reachable sets for linear dynamical systems*. Information and Control, vol. 19, no. 4, pages 319 – 344, 1971. (Cited on page 45.)

[Pfleeger *et al.* 2010] S. L. Pfleeger, J. B. Predd, J. Hunker and C. Bulford. *Insiders Behaving Badly: Addressing Bad Actors and Their Actions*. IEEE Transactions on Information Forensics and Security, vol. 5, no. 1, pages 169–179, 2010. (Cited on page 16.)

[Pivarčiová & Qazizada 2018] E. Pivarčiová and M. E. Qazizada. *Applications of process control system with (SCADA) and PID controller*. In 2018 ELEKTRO, pages 1–6, 2018. (Cited on page 21.)

[Ponomarev & Atkison 2016] S. Ponomarev and T. Atkison. *Industrial Control System Network Intrusion Detection by Telemetry Analysis*. IEEE Transactions on Dependable and Secure Computing, vol. 13, no. 2, pages 252–260, March 2016. (Cited on page 32.)

[Prada 2013] Cesar de Prada. Overview: Control hierarchy of large processing plants, pages 1–9. Springer London, London, 2013. (Cited on page 23.)

[Proctor & Zandt 2018] R. Proctor and T. Zandt. Human factors in simple and complex systems. CRC Press, 3rd édition, 2018. (Cited on page 12.)

[Puglia 2012] Luiz Vasco Puglia. *LINEAR PROGRAMMING FOR THE OPTIMAL CONTROL OF A ONE-DEGREE-OF-FREEDOM OVERHEAD CRANE SYSTEM*. 2012. (Cited on page 45.)

[Robla-Gómez *et al.* 2017] S. Robla-Gómez, V. M. Becerra, J. R. Llata, E. González-Sarabia, C. Torre-Ferrero and J. Pérez-Oria. *Working Together: A Review on Safe Human-Robot Collaboration in Industrial Environments*. IEEE Access, vol. 5, pages 26754–26773, 2017. (Cited on page 8.)

[Rrushi & Kang 2009] Julian Rrushi and Kyoung-Don Kang. *Detecting Anomalies in Process Control Networks*. In Charles Palmer and Sujeet Shenoi, editors, Critical Infrastructure Protection III, pages 151–165. Springer Berlin Heidelberg, March 2009. (Cited on page 32.)

[Rushby 1994] John Rushby. *Critical system properties: survey and taxonomy*. Reliab. Eng. Syst. Saf., vol. 43, no. 2, pages 189 – 219, 1994. (Cited on page 8.)

[Samtani *et al.* 2016] S. Samtani, S. Yu, H. Zhu, M. Patton and H. Chen. *Identifying SCADA vulnerabilities using passive and active vulnerability assessment techniques*. In 2016 IEEE Conference on Intelligence and Security Informatics (ISI), pages 25–30, 2016. (Cited on page 23.)

[Sauter 2010] T. Sauter. *The Three Generations of Field-Level Networks—Evolution and Compatibility Issues*. IEEE Transactions on Industrial Electronics, vol. 57, no. 11, pages 3585–3595, 2010. (Cited on page 14.)

[Sciascera *et al.* 2016] C. Sciascera, M. Galea, P. Giangrande and C. Gerada. *Lifetime consumption and degradation analysis of the winding insulation of electrical machines*. In 8th IET International Conference on Power Electronics, Machines and Drives (PEMD 2016), pages 1–5, 2016. (Cited on page 42.)

[Shapiro *et al.* 2011] Rebecca Shapiro, Sergey Bratus, Edmond Rogers and Sean Smith. *Identifying Vulnerabilities in SCADA Systems via Fuzz-Testing*. In Jonathan Butts and Sujeet Shenoi, editors, Critical Infrastructure Protection V, pages 57–72, Berlin, Heidelberg, 2011. (Cited on page 23.)

[Sicard *et al.* 2018] F. Sicard, C. Escudero, E. Zamaï and J.-M. Flaus. *From ICS attacks' analysis to the S.A.F.E. approach: implementation of filters based on behavioral models and critical state distance for ICS cybersecurity*. In Proc. CSNet, pages 1–8, Paris, October 2018. (Cited on pages 15 and 28.)

[Sicard *et al.* 2019] Franck Sicard, Éric Zamaï and Jean-Marie Flaus. *An approach based on behavioral models and critical states distance notion for improving cybersecurity of industrial control systems*. Reliability Engineering & System Safety, March 2019. (Cited on page 33.)

[Sicard 2018] Franck Sicard. *Prise en compte des risques de cyber-attaques dans le domaine de la sécurité des systèmes cyber-physiques : proposition de mécanismes*

de détection à base de modèles comportementaux. Theses, Université Grenoble Alpes, October 2018. (Cited on page 32.)

[Sikanen *et al.* 2018] Eerik Sikanen, Janne Nerg, Janne E. Heikkinen, Mohammad Gerami Tehrani and Jussi Sopanen. *Fatigue life calculation procedure for the rotor of an embedded magnet traction motor taking into account thermomechanical loads.* Mech. Syst. Signal Pr., vol. 111, pages 36 – 46, Oct. 2018. (Cited on pages 99 and 105.)

[Sridhar & Govindarasu 2014] Siddharth Sridhar and Manimaran Govindarasu. *Model-Based Attack Detection and Mitigation for Automatic Generation Control.* IEEE Transactions on Smart Grid, vol. 5, no. 2, pages 580–591, March 2014. (Cited on page 33.)

[Steusloff 1984] H. U. Steusloff. *Advanced Real-Time Languages for Distributed Industrial Process Control.* Computer, vol. 17, no. 2, page 37–46, Feb. 1984. (Cited on page 20.)

[Stouffer *et al.* 2015] K. Stouffer, V. Pillitteri, S. Lightman, M. Abrams and A. Hahn. *Guide to Industrial Control Systems (ICS) Security.* Technical report, National Institute of Standards and Technology, May 2015. (Cited on page 8.)

[Syrbe 1978] M. Syrbe. *Basic Principles of Advanced Process Control System Structures and a Realization with Distributed Microcomputers.* IFAC Proceedings Volumes, vol. 11, no. 1, pages 393 – 401, 1978. 7th Triennial World Congress of the IFAC on A Link Between Science and Applications of Automatic Control, Helsinki, Finland, 12-16 June. (Cited on page 20.)

[Štefan Kozák 2014] Štefan Kozák. *State-of-the-art in control engineering.* Journal of Electrical Systems and Information Technology, vol. 1, no. 1, pages 1 – 9, 2014. (Cited on page 21.)

[Teixeira 2019] A. M. H. Teixeira. *Optimal stealthy attacks on actuators for strictly proper systems.* In Proc. CDC, pages 4385–4390, Nice, Dec. 2019. (Cited on page 46.)

[Tendjaoui *et al.* 1991] Mustapha Tendjaoui, Christophe Kolski and Patrick Millot. *An approach towards the design of intelligent man-machine interfaces used in process control.* International Journal of Industrial Ergonomics, vol. 8, no. 4, pages 345 – 361, 1991. (Cited on page 23.)

[Thomesse 2005] J. P. Thomesse. *Fieldbus Technology in Industrial Automation.* Proceedings of the IEEE, vol. 93, no. 6, pages 1073–1101, 2005. (Cited on page 13.)

[Thuillier *et al.* 2020] Julien Thuillier, David Delouche, Jacques Fantini and Frédéric Kratz. *Detection of integrity loss in networked control systems using an interval finite memory observer.* Int. J. Control, vol. 0, no. 0, pages 1–10, Feb. 2020. (Cited on page 33.)

[Tulsyan & I. Barton 2016] Aditya Tulsyan and Paul I. Barton. *Reachability-based fault detection method for uncertain chemical flow reactors.* IFAC-PapersOnLine, vol. 49, no. 7, pages 1 – 6, 2016. 11th IFAC Symposium on Dynamics and Control of Process SystemsIncluding Biosystems DYCOPS-CAB 2016. (Cited on page 45.)

[Umsonst & Sandberg 2018] David Umsonst and Henrik Sandberg. *Anomaly Detector Metrics for Sensor Data Attacks in Control Systems.* In Proc. ACC, pages 153–158, Milwaukee, WI, June 2018. (Cited on page 39.)

[Umsonst *et al.* 2017] David Umsonst, Henrik Sandberg and Alvaro A. Cardenas. *Security analysis of control system anomaly detectors.* In Proc. ACC, pages 5500–5506, Seattle, WA, May 2017. (Cited on page 39.)

[Upadhyay *et al.* 2020] Darshana Upadhyay, Srinivas Sampalli and Bernard Plourde. *Vulnerabilities' Assessment and Mitigation Strategies for the Small Linux Server, Onion Omega2.* Electronics, vol. 9, no. 6, 2020. (Cited on page 23.)

[Valentine & Farkas 2011] S. Valentine and C. Farkas. *Software security: Application-level vulnerabilities in SCADA systems.* In 2011 IEEE International Conference on Information Reuse Integration, pages 498–499, 2011. (Cited on page 22.)

[VanAntwerp & Braatz 2000] Jeremy G. VanAntwerp and Richard D. Braatz. *A tutorial on linear and bilinear matrix inequalities.* Journal of Process Control, vol. 10, no. 4, pages 363 – 385, 2000. (Cited on page 53.)

[Venkatasubramanian *et al.* 2003] Venkat Venkatasubramanian, Raghunathan Rengaswamy, Surya N. Kavuri and Kewen Yin. *A review of process fault detection and diagnosis: Part III: Process history based methods.* Computers & Chemical Engineering, vol. 27, no. 3, pages 327 – 346, 2003. (Cited on page 32.)

[Villemeur 1992] A. Villemeur. Reliability, availability, maintainability and safety assessment, assessment, hardware, software and human factors. Wiley, 1992. (Cited on page 11.)

[Vinod *et al.* 2009] P. Vinod, V. Laxmi and M. S. Gaur. *Survey on Malware Detection Methods.* In Workshop on Computer and Internet Security, page 73–79, 2009. (Cited on page 32.)

[Vitturi *et al.* 2019] S. Vitturi, C. Zunino and T. Sauter. *Industrial Communication Systems and Their Future Challenges: Next-Generation Ethernet, IIoT, and 5G.* Proceedings of the IEEE, vol. 107, no. 6, pages 944–961, 2019. (Cited on page 14.)

[Viveros *et al.* 2014] P. Viveros, E. Zio, C. Nikulin, R. Stegmaier and G. Bravo. *Resolving equipment failure causes by root cause analysis and theory of inventive problem solving.* Proc Inst Mech Eng O J Risk Reliab, vol. 228, no. 1, pages 93–111, 2014. (Cited on page 12.)

[Vrignat *et al.* 2015] P. Vrignat, M. Avila, F. Duculty and F. Kratz. *Failure Event Prediction Using Hidden Markov Model Approaches.* IEEE Transactions on Reliability, vol. 64, no. 3, pages 1038–1048, 2015. (Cited on page 12.)

[Vukic 2003] Z. Vukic. Nonlinear control systems. Automation and Control Engineering. Taylor & Francis, 2003. (Cited on page 91.)

[Walker *et al.* 1981] D. N. Walker, S. L. Adams and R. J. Placek. *Torsional Vibration and Fatigue of Turbine-Generator shafts.* IEEE Transactions on Power Apparatus and Systems, vol. PAS-100, no. 11, pages 4373–4380, 1981. (Cited on page 42.)

[Wie 2008] B. Wie. Space vehicle dynamics and control, second edition. 2008. (Cited on page 105.)

[Williams 1990] T.J. Williams. *A Reference Model for Computer Integrated Manufacturing from the Viewpoint of Industrial Automation.* IFAC Proceedings Volumes, vol. 23, no. 8, Part 5, pages 281 – 291, 1990. 11th IFAC World Congress on Automatic Control, Tallinn, 1990 - Volume 5, Tallinn, Finland. (Cited on page 10.)

[Wucherer 2001] Klaus Wucherer. *HMI, The Window to the Manufacturing and Process Industry.* 8th IFAC Symposium on Analysis, Design and Evaluation of Human-Machine Systems (HMS 2001) (IFAC Proceedings Volumes), vol. 34, no. 16, pages 101 – 108, 2001. (Cited on page 23.)

[Yang *et al.* 2012] Liying Yang, Juntong Qi and Jianda Han. *Path planning methods for mobile robots with linear programming.* pages 641–646, 01 2012. (Cited on page 45.)

[Yaseen & Bayart 2017] Amer Atta Yaseen and Mireille Bayart. *Attack-tolerant networked control system: an approach for detection the controller stealthy hijacking attack.* Journal of Physics: Conference Series, vol. 783, page 012022, jan 2017. (Cited on page 33.)

[Ylmaz *et al.* 2018] E. N. Ylmaz, B. Ciylan, S. Gönen, E. Sindiren and G. Karacayılmaz. *Cyber security in industrial control systems: Analysis of DoS attacks against PLCs and the insider effect.* In 2018 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG), pages 81–85, 2018. (Cited on page 15.)

[Zamai *et al.* 1998] E. Zamai, A. Chaillet-Subias and M. Combacau. *An architecture for control and monitoring of discrete events systems.* Computers in Industry, vol. 36, no. 1, pages 95 – 100, 1998. (Cited on page 8.)

[Zhang & Venkitasubramaniam 2017] R. Zhang and P. Venkitasubramaniam. *Stealthy Control Signal Attacks in Linear Quadratic Gaussian Control Systems: Detectability Reward Tradeoff.* IEEE Trans. Inf. Forensics Security, vol. 12, no. 7, pages 1555–1570, Jul. 2017. (Cited on page 46.)

[Zhang *et al.* 2016] Heng Zhang, Peng Cheng, Ling Shi and Jiming Chen. *Optimal DoS Attack Scheduling in Wireless Networked Control System.* IEEE Trans. Control Syst. Technol., vol. 24, no. 3, pages 843–852, May 2016. (Cited on page 38.)

[Zhang *et al.* 2018] H. Zhang, Y. Qi, J. Wu, L. Fu and L. He. *DoS Attack Energy Management Against Remote State Estimation.* IEEE Control Netw. Syst., vol. 5, no. 1, pages 383–394, March 2018. (Cited on page 38.)

[Zhang *et al.* 2019a] Heng Zhang, Wenchao Meng, Junjian Qi, Xiaoyu Wang and Wei Xing Zheng. *Distributed Load Sharing Under False Data Injection Attack in an Inverter-Based Microgrid.* IEEE Trans. Ind. Informat., vol. 66, no. 2, pages 1543–1551, February 2019. (Cited on page 39.)

[Zhang *et al.* 2019b] Q. Zhang, K. Liu, Y. Xia and A. Ma. *Optimal Stealthy Deception Attack Against Cyber-Physical Systems.* IEEE Trans. Cybern., pages 1–10, 2019. (Cited on page 40.)

[Zhou & Baras 2015] Y. Zhou and J. S. Baras. *Reachable set approach to collision avoidance for UAVs.* In 2015 54th IEEE Conference on Decision and Control (CDC), pages 5947–5952, 2015. (Cited on page 45.)

[Zolotová & Landryová 2000] Iveta Zolotová and Lenka Landryová. *SCADA/HMI Systems and Emerging Technologies.* IFAC Workshop on Programmable Devices and Systems (PDS 2000) (IFAC Proceedings Volumes), vol. 33, no. 1, pages 17 – 20, 2000. (Cited on page 23.)

Abstract: Industrial Control Systems (ICSs) are architectures controlling a physical system to achieve an industrial objective. They are present in various sectors including energy generation and distribution, water treatment, manufacturing production, aerospace and defense. Those architectures are equipped with various digital entities (e.g. HMI, PLC, regulator) organized for the decrease in complexity of the control. Historically, ICSs have been designed to improve the productivity, but the cybersecurity has not been considered. Due to this lack, ICSs are facing cyberattacks. Plenty of them manipulating the architecture have been reported in the literature. They aim to steal sensitive information or to violate the integrity of the physical system. The violation of the physical system integrity refers to an intended alteration or destruction of the physical system through its control. It leads to a partial or complete failure of the services delivered by the physical system. Hence, cyberattacks are a new root-cause of failure, that we call the malicious acts. Those malicious acts aim to create and propagate anomalies in the architecture by exploiting vulnerabilities of the digital entities.

The research works developed in this thesis contribute in the prevention of the malicious acts in the ICS architectures. In particular, the contributions focus on the prevention of attacks targeting the controllers, of type of industrial regulator (e.g. PID controller). The type of attacks we address aim to accelerate the aging of the actuators through the control signal. Thus, this thesis considers this type of attacks by assuming an attacker who can modify the control signal.

About the contributions, they can be gathered in two propositions. Firstly, we propose to analyze those attacks. A first method of analysis aims to demonstrate the feasibility of such attacks by designing the control signal an attacker could inject to accelerate the aging of the actuators. It consists in solving an optimization problem with Sequential Linear Programming (SLP). A second method of analysis, based on a set theoretic approach, is developed to prove the feasibility of such attacks. It consists in quantifying the effect of those attacks on the physical system. Secondly, we propose to prevent the occurrence of those attacks. This method is derived from the previous one based on a set theoretic approach. It consists in restraining the control signal to make those attacks unfeasible. The main idea of the set theoretic-based methods is to search for a pseudo Lyapunov function. Two cases are considered: the quantification of the effects on an infinite-time horizon based on the concepts of positive invariance, and the quantification of the effects on a finite time horizon based on techniques from the robust convex simulation. In both cases, the methods consist in solving an optimization problem with semidefinite programming (SDP) under constraints expressed in terms of Linear Matrix Inequalities (LMI). Lastly, the proposed methods are validated on experiments.

Keywords: Stealthy attacks, Cybersecurity, Industrial Control Systems, Set-theoretic methods, Reachability analysis, Convex optimization

Résumé: Les systèmes de contrôle-commande industriels (ICSs) sont des architectures permettant de piloter un système physique afin de satisfaire un objectif industriel fixé. Ils sont présents dans de nombreux secteurs tels que la production et la distribution d'énergie, le traitement des eaux usés, le manufacturier, l'aérospatial ou encore la défense. Ces architectures ICSs s'appuient sur de nombreuses entités numériques (IHM, API, régulateurs, etc.) organisées pour maîtriser la complexité du pilotage du système physique considéré. Historiquement, ces ICSs ont été conçus afin d'améliorer la productivité, mais sans considérer nativement la cybersécurité. Cette absence de prise en compte conduit aujourd'hui les ICSs à être exposés à des cyberattaques ; nombre d'entre-elles manipulant le système de commande ont été révélées dans la littérature. Parmi ces attaques, certaines visent le vol de données confidentielles, d'autres visent à violer l'intégrité du système physique ; on parlera alors d'altération ou de destruction intentionnelle du système physique au travers de sa commande, entrainant de fait un dysfonctionnement partiel ou total des services rendus par le système physique. Les cyberattaques constituent ainsi de nouvelles causes racines de défaillances ; nous parlerons d'actes de malveillance. Ces actes de malveillance exploitent les vulnérabilités des entités numériques afin de créer et de propager des anomalies au sein de l'architecture ICS.

Les travaux développés dans cette thèse se proposent d'apporter leurs contributions dans le domaine de la prévention d'actes de malveillance envers les architectures de contrôle-commande industriel. Ils se focalisent en particulier sur la proposition d'une approche de prévention d'attaques ciblant les contrôleurs de type régulateurs industriels. Le type d'attaque adressé par ses travaux concerne celles qui accélèrent le vieillissement des actionneurs au travers du signal de commande. Cette thèse se place ainsi dans le cadre de ce type d'attaque en considérant un attaquant capable de modifier le signal de commande.

Sur le plan des contributions, les travaux développés s'articulent autour de deux propositions. Dans un premier temps, nous proposons d'analyser ces attaques. Une première méthode d'analyse vise à démonter leur faisabilité par la construction d'un signal de commande malveillant à injecter pour accélérer le vieillissement. Elle consiste à résoudre un problème d'optimisation exprimée sous forme de programmation linéaire séquentielle (SLP). Une seconde méthode d'analyse, basée sur une approche ensembliste, est développée quant à elle pour prouver la faisabilité de ces attaques. Elle consiste à quantifier l'effet de ces attaques sur le système physique. Dans un second temps, nous proposons d'empêcher l'occurrence de ces attaques. La méthode reprend la précédente sur la quantification de l'effet des attaques sur le système physique. Elle consiste à restreindre le signal de commande afin de les rendre infaisables. L'idée générale de ces méthodes basées sur une approche ensembliste consiste en la recherche d'une pseudo-fonction de Lyapunov. Deux cas sont considérés: la quantification des effets sur un horizon de temps infini à partir des concepts d'invariance positive, et la quantification des effets sur un horizon de temps fini à partir de techniques basées sur la simulation robuste convexe. Dans ces deux cas, les méthodes consistent à résoudre un problème

d'optimisation Semi-Définie Positive (SDP) avec des contraintes formulées en inégalités matricielles linéaires (LMI). Enfin, des expérimentations sont proposées afin de valider l'ensemble des propositions théoriques proposées.

Mots clefs: Attaques cachées, Cybersécurité, Système de contrôle-commande industriel, Méthodes ensemblistes, Analyse d'atteignabilité, Optimisation convexe