Applications of trajectory-based analysis in optimization and control

Sharifi K., Arman

DOI
10.4233/uuid:42cf6e19-7746-4d7e-bb6c-b8c357aacaaf

Publication date
2019

Document Version
Final published version

Citation (APA)
Sharifi K., A. (2019). *Applications of trajectory-based analysis in optimization and control*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:42cf6e19-7746-4d7e-bb6c-b8c357aacaaf

Applications of Trajectory-Based Analysis

in Optimization and Control

APPLICATIONS OF TRAJECTORY-BASED ANALYSIS

IN OPTIMIZATION AND CONTROL

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus Prof.dr.ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates,
to be defended publicly on Tuesday 12 March 2019 at 12:30 p.m.

by

Arman SHARIFI KOLARIJANI

Master of Science in Systems and Control,
Delft University of Technology, Delft, the Netherlands,
born in Ghaemshahr, Iran.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	Chairperson
Dr. T. Keviczky,	Delft University of Technology, promotor
Dr. P. Mohajerin Esfahani,	Delft University of Technology, copromotor
Independent members:	
Prof. dr. ir. C. Vuik,	Delft University of Technology
Prof. dr. ir. A. van Keulen,	Delft University of Technology
Prof. dr. A. Jadbabaie,	Massachusetts Institute of Technology
Prof. dr. M. Johansson,	KTH Royal Institute of Technology

This dissertation has been completed in partial fulfillment of the requirements of the dutch institute of systems and control (disc) for graduate study.

ISBN 978-94-6384-018-7

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

To my family, Noam, and Norman

CONTENTS

SUMMARY

The synergy between optimization and control is a long standing tradition. In fact, this synergy is becoming more and more apparent because of the multi-disciplinary character of the most pressing, current engineering problems along with constant developments of these two fields. Historically, optimization methods have helped the control community to achieve their design goals formalized in some sort of objective function. On the other hand, control theory has provided a setting to interpret complicated aspects of optimization algorithms. In this thesis, we address three problem instances that lie on the boundary of optimization and control. We employ tools from one field to address a problem in the other field. Fundamentally, our proposed methods share a similar character: their analysis techniques are trajectory-based. In simple words, our proposed methods exploit the trajectories generated by the dynamics that represent each problem instance.

The first problem focuses on a 2nd-order, damped differential equation (ODE). This ODE along with its numerous variations have been used to develop or analyze various optimization algorithms, known as fast methods. As an alternative to the existing methods, we first amend the underling ODE with two types of state-dependent inputs, and then extend the resulting controlled dynamics to two hybrid control systems. Employing a trajectory-based analysis, both control laws are constructed to guarantee exponential convergence in a suboptimality measure. To show that the trajectories generated by each hybrid control system are well-posed, we demonstrate Zeno-freeness of solution trajectories in both cases. Furthermore, we propose a mechanism to determine a time-discretization step-size such that the resulting discrete-time hybrid control systems are exponentially stable.

Event-based implementation of control laws have received a lot of attention during the past decade. The reason for this interest is the hope to reduce the conservatism involved in the traditional periodic implementation. In the second problem of this thesis, we introduce an event-based sampling policy for a constraint-tightening, robust model predictive control (RMPC) method. The triggering mechanism is a sequence of hyper-rectangles constructed around the optimal state trajectories. In particular, the triggering mechanism's nature makes the proposed approach a suitable choice for plants without a centralized sensory node. A key feature of the proposed method is its complete decoupling from the RMPC method's parameters, facilitating a meaningful comparison between the periodic and aperiodic implementation policies. Furthermore, we provide two types of convex formulations to design the triggering mechanism.

The last problem we focus on in this thesis is also related to the event-based implementation of a control law. However, the main aim here is to propose an entity that can be utilized by a real-time engineer to schedule tasks in a networked structure. A common entity provided in the literature related to event-triggering approaches is the minimal inter-execution time (to show the avoidance of a Zeno behavior in the closed-loop sys-

tem). Nonetheless, such a quantity is extremely conservative when used for scheduling purposes. In this problem, we consider an $\mathscr{L}_2$-based triggering mechanism introduced in the literature and propose a framework to construct a timed safety automaton that can capture the triggering instants generated by this mechanism. In our analysis, we borrow some tools from stability analysis of delayed systems along with reachability analysis to construct the desired timed safety automaton.

SAMENVATTING

De synergie tussen optimalisatie en regeltechniek is er een met een lange traditie. Sterker nog, deze synergie is onderweg om nog meer zichtbaar te worden door het multidisciplinaire karakter van de meest uitdagende van de huidige technische problemen, naast de voortschrijdende ontwikkeling van deze twee vakgebieden. Historisch gezien hebben methodes uit de wereld van de optimalisatie regeltechnici geholpen om hun doelen, die vaak geformaliseerd in de vorm van een criteriumfunctie, te bereiken. Aan de andere kant hebben theorieën uit de regeltechniek geholpen om ingewikkelde aspecten van optimalisatiealgoritmen te interpreteren. In deze these worden drie gevallen behandeld die zich ergens op het snijvlak tussen optimalisatie en regeltechniek bevinden. Daarbij worden handvatten uit het ene vakgebied gebruikt om problemen in het andere veld aan te pakken. In de basis delen de in deze these voorgestelde methoden een gemeenschappelijk karakter: de analysetechnieken die gebruikt worden zijn allen trajectorie-gebaseerd. Samengevat: de voorgestelde methoden halen voordeel uit kennis van trajectories, die voortkomt uit de dynamische vergelijkingen die het betreffende probleemgeval representeren.

In het eerste probleem dat behandeld wordt richten we ons op een tweede-orde, gedempte 'gewone differentiaalvergelijking' (ODE). Deze differentiaalvergelijking, samen met talrijke varianten hierop, zijn gebruikt om diverse optimalisatiealgoritmen - te weten de zogenaamde 'fast methods' - te ontwikkelen en analyseren. Anders dan voor de tot nu toe bestaande methoden, voegen we eerst aan de differentiaalvergelijking twee typen toestandsbepaalde ingangssignalen toe, waarna we de resulterende dynamica - met de toegevoegde terugkoppelingssignalen - uitbreiden tot hybride regelsystemen. Deze beide regelschema's worden geconstrueerd, gebruikmakende van trajectorie-gebaseerde analyse, zodanig dat exponentiële convergentie gegarandeerd wordt ten opzichte van een bepaalde suboptimaliteitsmaatstaf. Om te laten zien dat de trajectories die door beide hybride regelsystemen welgesteld zijn wordt voor beide gevallen aangetoond dat er geen zogenaamd Zeno-gedrag voorkomt in de oplossingstrajectories. Daarbovenop wordt een mechanisme voorgesteld om een stapgrootte voor de tijds-discretisatie te bepalen zodanig dat de resulterende discrete-tijd hybride regelsystemen exponentieel stabiel zijn.

Gebeurtenisgebaseerde (event-based) implementaties van regelmechanismen hebben het laatste decennium veel aandacht gegenereerd. Deze aandacht wordt getrokken door de wens om het conservatisme in de reguliere periodieke implementatie van regelmechanismen te verminderen. In het tweede probleemgeval in dat in deze these besproken wordt introduceren we een gebeurtenisgebaseerd bemonsteringsbeleid voor een robuust modelgebaseerd voorspellend regelsysteem (RMPC), waarbij de randvoorwaarden verkrapt worden om robuustheid te garanderen (constraint tightening). Het activeringsmechanisme voor het opnieuw bemonsteren bestaat uit een reeks van hyperrechthoeken die geconstrueerd zijn rondom de optimale toestandstrajectorie. In het bij-

zonder maakt deze hyperrechthoekige natuur van het activeringsmechanisme het mogelijk om deze toe te passen op systemen zonder gecentraliseerd sensorknooppunt. Een belangrijk kenmerk van de voorgestelde methode is dat deze compleet ontkoppeld is van de parameters van het voorspellend regelsysteem, waardoor een betekenisvolle vergelijking tussen het periodieke en gebeurtenisgebaseerde implementatiebeleid gemaakt kan worden. Er worden voor deze methode twee typen convexe formuleringen voor het bepalen van de hyperrechthoeken aangereikt.

Het laatste probleem waar we ons in deze these op richten is ook verbonden aan het gebeurtenisgebaseerd implementeren van een regelmechanisme. Anders dan hiervoor het geval ligt de focus hier op het voorstellen van een entiteit die gebruikt kan worden door een realtime-ingenieur om taken in een netwerkstructuur in te plannen. Een entiteit die in de literatuur vaak gevonden wordt in relatie tot gebeurtenisgeactiveerde aanpakken is de minimale tijd tussen twee uitvoeringen van een taak (minimal interexecution time), om te laten zien dat in het gesloten-lus systeem geen Zeno-gedrag voorkomt. Daarentegen, zo'n grootheid is vaak extreem conservatief wanneer deze gebruikt wordt voor planningstoepassingen. In dit laatste probleemgeval beschouwen we een L2-gebaseerd activeringsmechanisme die in de literatuur geïntroduceerd is en stellen we een kader voor om een afgeklokt veiligheidsautomaton te construeren. Voor dit automaton is het mogelijk om de activeringstijdstippen van het mechanisme vast te leggen. In onze analyse lenen we enkele instrumenten uit stabiliteitsanalyse van systemen met vertragingen, samen met instrumenten uit de haalbaarheidsanalyse, om het gewenste geklokte veiligheidsautomaton te construeren.

ACKNOWLEDGEMENTS

My deepest gratitude goes to my promotor Tamás Keviczky. His constant support and encouragement along with his willingness and trust in allowing me to seek problems of my interest made my whole PhD studies an enjoyable experience. Many thanks also go to my copromotor Peyman Mohajerin Esfahani. His attitudes, such as the effort to fundamentally comprehend a subject and relentless simplifications of a result in order to make a theoretical statement accessible to a greater audience, were constant sources of my surprise during our collaborations. I really appreciate Bart De Schutter for being my promotor for the first two years of my PhD studies. I am also grateful to Manuel Mazo Jr., who was my Master thesis supervisor. Besides the fact that one of the chapters of this thesis is a result of the collaboration with him, his problem solving attitude has greatly impacted mine. I would like to express my sincere gratitude to the defense committee members Fred van Keulen, Kees Vuik, Ali Jadbabaie and Mikael Johansson for their constructive criticism and feedback that improved the quality of this text. Some part of this thesis is a result of collaborations with my dear colleagues and friends, Sander Bregman and Dieky Adzkiya. It has been a great pleasure working with these two individuals. Moreover, I have to thank my great officemates Bart, Thomas, and Amin. Finally, I would like to thank my family members for their unlimited support.

Arman Sharifi Kolarijani
Delft, January 2019

1

INTRODUCTION

Questions related to dynamical systems represent a prominent element in many scientific and technological fields. This prominence stems from the fact that our ability to comprehend any phenomenon starts with a sensible object, the model. With regards to a phenomenon, the model is in fact the entity that allows one to construct logical statements, to interpret a specific behavior, or (if possible) to manipulate a certain aspect of the phenomenon via the environment that the phenomenon is in interaction with. In particular, this perspective enables us to categorize seemingly different phenomena from diverse scientific disciplines in a unified manner. By virtue of this mindset, intuitive conclusions from a certain discipline are carried over to another discipline. One is then capable of interpreting difficult to grasp behaviors in a rather straightforward manner.

Generally speaking, the specific behavior of a phenomenon that one wishes to analyze or to regulate is some sort of a functional of dynamics (i.e., the states) and/or inputs (if the phenomenon is excited by its surrounding environment). We shall call this specific behavior, the performance measure. A principal class of analysis techniques in dynamical systems is the so-called Lyapunov-based methodologies. These methodologies provide a simple universal approach to quantitatively study a performance measure. The key properties that make these methodologies such an omnipotent tool are their natural simplicity and physical intuitions. They are undeniably a corner-stone in the analysis of dynamical systems. The evidence is their numerous extensions and their versatility in a variety of fields, such as optimization, statistical analyses, verification, identification, etc.

Despite the universality of Lyapunov-based methodologies, it is probable that the application of such a generic notion will limit one's ability to exploit subtle features of a certain problem instance. This matter then becomes crucial when one seeks the best possible outcome for the considered performance measure of a phenomenon. Let us now elaborate on the above statement by providing an example. Consider an iterative optimization algorithm. It is not difficult to see that the algorithm can be understood as a discrete-time dynamical system $x_{k+1} = f(k, x_k)$, where k denotes the iteration index, $x_k \in \mathbb{R}^n$ is the decision variable, and the map $f : \mathbb{Z}_{\geq 0} \times \mathbb{R}^n \to \mathbb{R}^n$ represents the dynam-

ical system capturing the algorithm's dynamics. Let us further call the best achievable outcome for a certain performance measure of an algorithm, an "optimal" outcome. Roughly speaking, most of the approaches in the literature that provide optimal performance measures (e.g., iteration complexity, rate of convergence, etc.) heavily exploit the trajectories generated by their corresponding map f. (Notice that there are some "optimal" methods in the literature that do possess equivalent Lyapunov-based reformulations.) The difficulty of finding a Lyapunov function for a general nonlinear dynamical system is a known fact in control theory. Most of the proposed trajectory-based algorithms heavily depend on the problem instance that these algorithms deal with. As a result, if not impossible, it requires a lot of effort to extend these results to other settings while using sophisticated argumentation.

In this thesis we argue in the favor of such specifically tailored tools related to three problem instances that are clearly different on the surface, namely, (i) a dynamical system viewpoint to synthesize a fast optimization algorithm, (ii) an event-based policy to implement a robust model predictive control approach, and (iii) a symbolic framework that captures the timing behavior of perturbed linear time-invariant systems with an event-triggered implementation. The three proposed approaches follow a fundamentally similar concept: they exploit the underlying structure of the solution trajectories generated by each problem instance. In what follows, we introduce these three problems and outline the proposed approaches to solve each problem, at a conceptual level.

Chapter 2: This chapter deals with the continuous-time counterpart of a class of iterative optimization algorithms, the so-called fast gradient-based methods, as the underlying phenomenon. (The prefix fast refers to the fact that these methods have an order of magnitude higher rate of convergence compared to non-fast methods.) The dynamical system that models the counterpart is a generic, damped 2nd-order ordinary differential equation (ODE).

The treatment of optimization methods as dynamical systems can be traced back centuries ago. The motivations behind such a treatment relies on the following fact: the ability to utilize tools from dynamical systems whether to interpret behaviors of existing optimization methods or to design new optimization methods. Recently, it has been observed that if the step size of Nesterov's celebrated fast method is pushed toward zero, his algorithm recovers a damped, 2nd-order ODE with a time-dependent damping term. Prior to Nesterov and surprisingly, Polyak also based his celebrated momentum method on a similar damped, 2nd-order ODE, instead with a constant damping term. The observation regarding Nesterov's fast method along with its relationship to Polyak's momentum method have become the driving force in the algorithmic optimization community to design new optimization methods inspired by damped, 2nd-order ODEs. Indirectly, this mindset has also restored the interest in the application of numerical discretizations not only on the damped, 2nd-order ODEs but also on gradient systems (which are 1st-order ODEs). In other words, researchers have become curious about the beneficiary impacts of more advanced temporal discretization methods in designing and/or interpreting optimization algorithms.

In the same vein as the recent dynamical system viewpoint, we propose two classes of fast methods that are formulated as hybrid control systems. We focus on an unconstrained, smooth problem $\min_{X\in\mathbb{R}^n} f(X)$, where $X \in \mathbb{R}^n$ is the decision variable, the func-

tion $f : \mathbb{R}^n \to \mathbb{R}$ is the objective function with f^* denoting the minimum of the function f. The proposed frameworks accept a user-defined positive scalar α and guarantee the exponential convergence rate $\mathcal{O}(e^{-\alpha t})$ in the suboptimality measure $f(X(t)) - f^*$. The reasoning behind extending the class of dynamical systems represented by ODEs to the ones represented by a mixture of continuous-discrete dynamics (i.e., the hybrid formulation) stems from approaches used in the literature to address an unwanted behavior of fast methods. Fast methods are not descent and to counteract such a behavior most of the approaches employ some sort of restarting schemes. Evidently, one can consider such optimization algorithms as dynamical systems with jumps. A natural choice of modeling is then a hybrid formulation. Alternative to the existing fast methods in which the damping term of the 2nd-order ODE is time-dependent, we dynamically synthesize feedback controls in a state-dependent manner. The input synthesis approaches are trajectory-based rather than Lyapunov-based. In the first proposed class, the damping term is viewed as the control input. In the second class, the amplitude with which the gradient of the objective function impacts the dynamics serves as the controller. Here, the damping term is constant. The objective function requires to satisfy a certain sharpness criterion, the so-called Polyak–Łojasiewicz inequality. Moreover, we establish that both hybrid structures possess Zeno-free solution trajectories. We finally provide a mechanism to determine the discretization step size to attain an exponential convergence rate. The materials presented in this chapter are previously reported in [1], [2], and [3].

Chapter 3: In this cahpter, the considered phenomenon is a perturbed linear time-invariant (LTI) system controlled with a certain robust model predictive control (RMPC) method. The performance measure, we seek to study, is the possibility of reducing the communication and/or computation loads of this control approach in a networked environment.

In recent trends of applications, multiple subsystems sometimes cooperatively and other times competitively utilize shared resources in a networked structure. Two main resources that are commonly present in such networked structures are communication and computation components. During the last two decades, increased attention has been devoted to such networked problems and the term “networked control systems" was coined to refer to these problems. In early studies, the beneficiary properties of implementing a control law in an aperiodic fashion, as opposed to a traditional periodic (time-triggered) fashion, have been pointed out. Subsequently, Tabuada demonstrated in his seminal paper that one can employ the principles of input-to-state stability to intelligently implement a control action in an aperiodic way while guaranteeing the closed-loop stability. As a result, one expects the possibility to reduce communication and computation burdens on shared resources. His mindset then spread out in many directions which can be identified by the term event-triggering control in the literature. (It is worth mentioning that the responsible entity that dictates the time to close a control loop is also known as the triggering mechanism.) Interestingly enough, this spread is not just limited to the problems that are of interest to control theory communities. For example, consider an optimization problem in which multiple agents try to cooperatively minimize a global objective function. In such a context, a similar idea has been employed in several methods in order to define a communication policy to exchange

local information while guaranteeing convergence to the global minimum.

In control methodologies, the class of RMPC methods is a potent candidate to be equipped with an event-based implementation. This statement is simply supported by the fact that the predicted trajectories at a given sampling instant can be considered as a basis to construct a triggering mechanism. In this chapter, an event-based sampling approach is proposed for a constraint-tightening RMPC method in the literature. The proposed approach is applicable to perturbed, LTI systems with polytopic input/state constraints. In a geometrical sense, the triggering mechanism is a sequence of hyper-rectangles constructed around the optimal state trajectory. We show that employing the proposed event-based implementation, robust recursive feasibility and robust stability are guaranteed. In particular, the triggering mechanism is applicable to plants without a centralized sensory component (that collects and is aware of all states). This is simply because the triggering mechanism can be evaluated locally at each individual sensor. The design of the triggering mechanism is cast as a constrained parametric-in-set optimization problem with the volume of set as the objective function. We establish that one can reformulate the optimization problem of triggering mechanism's design in terms of vertices of the desired hyper-rectangles. By doing so, the design problem becomes a convex nonlinear program. We further borrow some mathematical tools from the literature to show that a linear program reformulation of the triggering mechanism's design is also possible. The materials presented in this chapter are previously reported in [4] and [5].

Chapter 4: The phenomenon encountered in this chapter is the class of perturbed LTI systems controlled with an event-triggering implementation. The performance measure of interest is an object that formally captures the triggering instants generated by the triggering mechanism.

In networked control systems, the application of event-triggering strategies in the sampling process is expected to reduce the usage of network resources, such as communication bandwidth. However, it is essential to bring to the reader's attention the following two facts regarding event-triggering implementations. First, most of the studies in the literature fall short of guaranteeing improvements by employing such aperiodic implementations. The analysis provided by these studies usually lacks a mathematically sound framework to compare event- and time-triggering implementations. Nonetheless, numerical and experimental case studies support the claim that event-triggering approaches on average can outperform time-triggering counterparts for a certain behavior of the closed-loop system (e.g., the required average number of instants to sample the states). Another assertion regarding event-triggering implementations, that is commonly claimed without a rigorous proof, is the possibility of improving the energy consumption of communication components and/or the computational effort to derive the control input. On the surface, such a claim seems rather logical. However, this claim does not always hold on a practical level.

Let us clarify the previous statement regarding the energy consumption in a communication network. Generally speaking, two groups are involved in the design process of real-time systems, in particular networked control systems: control engineers who design control laws, and real-time engineers who design schedulers to safely implement tasks in the communication network. Suppose one considers a time-triggering implementation of the control law. The only piece of information that should be conveyed be-

tween these two groups is simply a periodic or an aperiodic time period based on which the control law is updated. Such a concept is referred to as the “separation of concerns" in the scheduling community. Consider now an event-triggering implementation of the control law. In this case, there is no such simple object (e.g., the time period) to be provided to real-time engineers for a proper task scheduling. This is due to the fact that the next triggering instant is now implicitly determined by the triggering mechanism. (Notice that most of the event-triggering mechanisms are state-dependent functions.) Thus, the aperiodic nature of sampling periods generated by event-triggering approaches hinder the schedulability of such networked control systems.

In this chapter, we propose a formal framework to construct a timed safety automaton that captures the sampling behavior of an event-triggering approach for perturbed LTI systems. The triggering mechanism is designed based on the $\mathscr{L}_2$ stability, introduced in the literature. The term sampling behavior refers to all sequences of triggering instants that can be generated by the triggering mechanism. The proposed framework consists of two main stages in order to construct the time safety automaton. In the first stage, the state-space is partitioned into a finite number of convex polyhedral cones. Each cone represents a discrete mode in the abstracted automaton. Inspired by an approach introduced in the literature, we then (i) adopt techniques from stability analysis of retarded systems and (ii) use a polytopic embedding of time to construct several LMI conditions. These LMI conditions enable us to characterize the sampling interval associated with each region. This time interval denotes all the possible sampling periods that can be generated by the triggering mechanism, assuming the state at the last triggering instant lies inside the corresponding cone. In the second stage, we use reachability analysis to derive all the transitions in the abstracted automaton. This leads to an object that can be passed to real-time engineers for the scheduling purpose. The materials of this chapter are previously reported in [6].

Chapter 5: In the final chapter, we conclude this thesis and provide several future research directions for each of the three topics.

2

CONTINUOUS-TIME ACCELERATED METHODS VIA A HYBRID CONTROL LENS

Inspired by the recent dynamical system viewpoint of Nesterov's fast method, we propose two classes of fast optimization methods in this chapter. We formulate the dynamics of these optimization methods in the framework of hybrid control systems. Alternative to the existing fast methods which are parametric-in-time second order differential equations, we synthesize the feedback inputs in a state-dependent fashion. In the first class, a damping term is viewed as the control input. In the second class, the amplitude with which the gradient of the objective function impacts the dynamics serves as the controller. The structures of the inputs are determined such that a pre-specified exponential convergence rate is guaranteed. The proposed methodologies require the objective function to satisfy a certain sharpness criterion, the so-called Polyak–Łojasiewicz inequality. We further establish that both of the hybrid formulations possess Zeno-free solution trajectories. In order to construct the discrete-time counterparts of the proposed continuous-time hybrid control systems (that are the iterative optimization algorithms), we finally provide a mechanism to determine the discretization step size such that an exponential convergence rate can be attained. The materials presented in this chapter are previously reported in [1], [2], and [3].

2.1. INTRODUCTION

There is a renewed surge of interest in gradient-based algorithms in many computational communities such as machine learning and data analysis. The following non-exhaustive list of references indicates typical application areas: clustering analysis [7], neuro-computing [8], statistical estimation [9], support vector machines [10], signal and image processing [11], and networked-constrained optimization [12]. This interest primarily stems from low computational and memory loads of these algorithms (making them exceptionally attractive in large-scale problems where the dimension of decision variables can be enormous). As a result, a deeper understating of how these algorithms function has become a focal point of many studies.

One research direction that has been recently revitalized is the application of ordinary differential equations (ODEs) to the analysis and design of optimization algorithms. Consider an iterative algorithm that can be viewed as a discrete dynamical system, with the scalar s as its step size. As s decreases, one can observe that the iterative algorithm in fact recovers a differential equation, e.g., in the case of gradient descent method applied to an unconstrained optimization problem $\min_{X\in\mathbb{R}^n} f(X)$, one can inspect that

$$X^{k+1} = X^k - s\nabla f(X^k) \;\rightsquigarrow\; \dot{X}(t) = -\nabla f\big(X(t)\big)$$

where $f:\mathbb{R}^n \to \mathbb{R}$ is a smooth function, X is the decision variable, $k \in \mathbb{Z}_{\geq 0}$ is the iteration index, and $t \in \mathbb{R}_{\geq 0}$ is the time. The main motivation behind this line of research has to do with well-established analysis tools in dynamical systems described by differential equations.

The slow rate of convergence of the gradient descent algorithm ($\mathcal{O}(\frac{1}{t})$ in continuous and $\mathcal{O}(\frac{1}{k})$ in discrete time), limits its application in large-scale problems. In order to address this shortcoming, many researchers resort to the following class of 2nd-order ODEs, which is also the focus of this study:

$$\ddot{X}(t) + \gamma(t)\dot{X}(t) + \nabla f\big(X(t)\big) = 0. \tag{2.1}$$

Increasing the order of the system dynamics interestingly helps improve the convergence rate of the corresponding algorithms to $\mathcal{O}(\frac{1}{k^2})$ in the discrete-time domain or to $\mathcal{O}(\frac{1}{t^2})$ in the continuous-time domain. Such methods are called *momentum*, *accelerated*, or *fast* gradient-based iterative algorithms in the literature. The time-dependent function $\gamma : \mathbb{R}_{\geq 0} \to \mathbb{R}_{>0}$ is a *damping* or a *viscosity* term, which has also been referred to as the *asymptotically vanishing viscosity* since $\lim_{t\to\infty} \gamma(t) = 0$ [13].

Chronological developments of fast algorithms: It is believed that the application of (2.1) to speed-up optimization algorithms is originated from [14] in which Polyak was inspired by a physical point of view (i.e., a heavy-ball moving in a potential field). Later on, Nesterov introduced his celebrated accelerated gradient method in [15] using the notion of "estimate sequences" and guaranteeing convergence rate of $\mathcal{O}(\frac{1}{k^2})$. Despite several extensions of Nesterov's method [16–18], the approach has not yet been fully understood. In this regard, many have tried to study the intrinsic properties of Nesterov's method such as [19–22]. Recently, the authors in [23] and in details [24] surprisingly discovered that Nesterov's method recovers (2.1) in its continuous limit, with the time-varying damping term $\gamma(t) = \frac{3}{t}$.

A dynamical systems perspective: Based on the observation suggested by [23], several novel fast algorithms have been developed. Inspired by the mirror descent approach [25], the ODE (2.1) has been extended to non-Euclidean settings and to higher order methods using the Bregman Lagrangian in [26]. Following [26], a "rate-matching" Lyapunov function is proposed in [27] with its monotonicity property established for both continuous and discrete dynamics. Recently, the authors in [22] make use of an interesting semidefinite programming framework developed by [21] and use tools from robust control theory to analyze the convergence rate of optimization algorithms. More specifically, the authors exploit the concept of integral quadratic constraints (IQCs) [28] to design iterative algorithms under the strong convexity assumption. Later, the authors in [29] extend the results of IQC-based approaches to quasi-convex functions. The authors in [30] use dissipativity theory [31] along with the IQC-based analysis to construct Lyapunov functions enabling rate analyses.

Restarting schemes: A characteristic feature of fast methods is the non-monotonicity in the suboptimality measure $f - f^*$, where f^* refers to the optimal value of function f. The reason behind such an undesirable behavior can be intuitively explained in two ways: (i) a momentum based argument indicating as the algorithm evolves, the algorithm's momentum gradually increases to a level that it causes an oscillatory behavior [32]; (ii) an acceleration-based argument indicating that the asymptotically vanishing damping term becomes so small that the algorithm's behavior drifts from an over-damped regime into an under-damped regime with an oscillatory behavior [24]. To prevent such an undesirable behavior in fast methods, an optimal fixed restart interval is determined in terms of the so-called condition number of function f such that the momentum term is restarted to a certain value, see e.g., [16, 18, 33–35]. It is worth mentioning that [32] proposes two heuristic adaptive restart schemes. It is numerically observed that such restart rules practically improve the convergence behavior of a fast algorithm.

Regularity for exponential convergence: Generally speaking, exponential convergence rate and the corresponding regularity requirements of the function f are two crucial metrics in fast methods. In what follows, we discuss about these metrics for three

popular fast methods in the literature. When the objective functions are strongly convex with a constant σ_f and their gradient is Lipschitz with a constant L_f, [24] proposes the "speed restarting" scheme

$$\sup\Big\{t>0:\ \forall \tau\in(0,t),\ \frac{d\|\dot{X}(\tau)\|^2}{d\tau}>0\Big\},$$

to achieve the convergence rate of:

$$f\big(X(t)\big)-f^*\leq d_1 e^{-d_2 t}\|X(0)-X^*\|^2.$$

The positive scalars d_1 and d_2 depend on the constants σ_f and L_f. Assuming the convexity of the function f with a certain choice of parameters in their "ideal scaling" condition, [26] guarantees the convergence rate of $\mathcal{O}(e^{-ct})$ for some positive scalar c. However, in this general case, their approach requires to compute a matrix inversion in the Euler-Lagrange equation in the form of:

$$\ddot{X}(t)+c\dot{X}(t)+c^2e^{ct}\Big(\nabla^2 h\big(X(t)+\frac{1}{c}\dot{X}(t)\big)\Big)^{-1}\nabla f\big(X(t)\big)=0,$$

where the function h is a distance generating function. Under uniform convexity assumption with a constant ν_f, it is further shown that

$$f\big(X(t)\big)-f^*\leq\Big(f\big(X(0)\big)-f^*\Big)e^{-\nu_f\frac{1}{p-1}t},$$

where $p-1$ is the order of smoothness of f. The authors in [27] introduce the Lyapunov function

$$\mathcal{E}(t)=e^{\beta(t)}\left(f\big(X(t)\big)-f^*+\frac{\sigma_f}{2}\|X^*-Z(t)\|^2\right),$$

to guarantee the rate of convergence

$$\mathcal{E}(t)\leq\mathcal{E}(0)e^{-\int\dot{\beta}(s)ds},$$

where $Z(t)=X(t)+\frac{1}{\dot{\beta}(t)}\dot{X}$, $\dot{Z}(t)=-\dot{X}(t)-\frac{1}{\sigma_f}\dot{\beta}(t)\nabla f\big(X(t)\big)$, and $\beta(t)$ is a user-defined function.

Contribution: Much of the references reviewed above primarily deal with constructing a time-dependent damping term $\gamma(t)$ that is sometimes tied to a Lyapunov function. Furthermore, due to underlying oscillatory behavior of the corresponding 2nd-order ODE, researchers utilize restarting schemes to over-write the steady-state non-monotonic regime with the transient monotonic regime of the dynamics. In general, notice that these schemes are based on time-dependent schedulers.

Statement of hypothesis: With the above argument in mind, let us view an algorithm as a unit point mass moving in a potential field caused by an objective function f under a parametric (or possibly constant) viscosity, similar to the second order ODE (2.1). In this view, we aim to address the following two questions: Is it possible to

I) synthesize the damping term γ as a state-dependent term (i.e., $\gamma(X,\dot{X})$), or

II) dynamically control the magnitude of the potential force $\nabla f(X)$,

such that the underlying properties of the optimization algorithm are improved?

In this chapter, we answer these questions by amending the 2nd-order ODE (2.1) in two ways as follows:

(I) $\ddot{X}(t) + u_{\mathbf{I}}\big(X(t), \dot{X}(t)\big)\, \dot{X}(t) + \nabla f(X(t)) = 0,$

(II) $\ddot{X}(t) + \dot{X}(t) + u_{\mathbf{II}}\big(X(t), \dot{X}(t)\big)\, \nabla f(X(t)) = 0,$

where the indices indicate to which question each structure is related to in the above hypothesis. Evidently, in the first structure, the state-dependent input $u_{\mathbf{I}}$ replaces the time-dependent damping γ in (2.1). While in the second structure, the feedback input $u_{\mathbf{II}}$ dynamically controls the magnitude with which the potential force enters the dynamics (we assume for simplicity of exposition that $\gamma(t) = 1$, however, one can modify our proposed framework and following a similar path develop the corresponding results for the case $\gamma(t) \neq 1$). Given a positive scalar α, we seek to achieve an exponential rate of convergence $\mathcal{O}(e^{-\alpha t})$ for an unconstrained, smooth optimization problem in the sub-optimality measure $f\big(X(t)\big) - f^*$. To do so, we construct the state-dependent feedback laws for each structure as follows:

$$
\begin{aligned}
u_{\mathbf{I}}\big(X(t), \dot{X}(t)\big) &:= \quad \alpha + \frac{\|\nabla f(X(t))\|^2 - \langle \nabla^2 f\big(X(t)\big)\dot{X}(t), \dot{X}(t)\rangle}{\langle \nabla f\big(X(t)\big), -\dot{X}(t)\rangle}, \\
u_{\mathbf{II}}\big(X(t), \dot{X}(t)\big) &:= \quad \frac{\langle \nabla^2 f\big(X(t)\big)\dot{X}(t), \dot{X}(t)\rangle + (1-\alpha)\langle \nabla f\big(X(t)\big), -\dot{X}(t)\rangle}{\|\nabla f(X(t))\|^2}.
\end{aligned}
$$

Motivated by restarting schemes, we further extend the class of dynamics to hybrid control systems (see Definition 2.2.1 for further details) in which both of the above ODE structures play the role of the *continuous flow* in their respective hybrid dynamical extension. We next suggest an admissible control input range $[u_{\min}, u_{\max}]$ that determines the *flow set* of each hybrid system. Based on the model parameters α, $u_{\min}$, and $u_{\max}$, we then construct the *jump map* of each hybrid control system by the mapping $\big(X^\top, -\beta\nabla^\top f(X)\big)^\top$ guaranteeing that the range space of the jump map is contained in its respective flow set. Notice that the velocity restart schemes take the form of $\dot{X} = -\beta\nabla f(X)$.

We now summarize the contributions of our proposed approaches in the context of continuous-time, fast methods:

- We introduce system theoretic frameworks to synthesize state-dependent feedback inputs given a prescribed control input bound and a desired convergence rate (Theorems 2.3.1 and 2.3.4). Notice that the state-dependent feature of our proposed dynamical systems differs from commonly time-dependent methodologies in the literature.

- We derive a lower bound on the time between two consecutive jumps for each hybrid structure. This ensures that the constructed hybrid systems admit the so-called Zeno-free solution trajectories. It is worth noting that the regularity assumptions required by the proposed structures are different (Theorems 2.3.2 and 2.3.5).

- The proposed frameworks are general enough to include a subclass of non-convex problems. Namely, the critical requirement is that the objective function f satisfies the Polyak–Łojasiewicz (PL) inequality (Assumption (A2)), which is a weaker regularity assumption than the strong convexity that is often assumed in this context.
- We utilize the *forward-Euler* method to discretize both hybrid systems (i.e., obtain optimization algorithms). We further provide a mechanism to compute the step size such that the corresponding discrete dynamics have an exponential rate of convergence (Theorem 2.3.11).

The remainder of this chapter is organized as follows. In Section 2.2, the mathematical notions are represented. The main results of the chapter are introduced in Section 2.3. Section 2.4 contains the proofs of the main results. We introduce a numerical example in Section 2.5. This chapter is finally concluded in Section 2.6.

Notations: The sets $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$ denote the n-dimensional Euclidean space and the space of $m\times n$ dimensional matrices with real entries, respectively. For a matrix $M\in\mathbb{R}^{m\times n}$, $M^\top$ is the transpose of M, $M\succ 0$ ($\prec 0$) refers to M positive (negative) definite, $M\succeq 0$ ($\preceq 0$) refers to M positive (negative) semi-definite, and $\lambda_{\max}(M)$ denotes the maximum eigenvalue of M. The $n\times n$ identity matrix is denoted by I_n. For a vector $v\in\mathbb{R}^n$ and $i\in\{1,\cdots,n\}$, v_i represents the i-th entry of v and $\|v\|:=\sqrt{\Sigma_{i=1}^n v_i^2}$ is the Euclidean 2-norm of v. For two vectors $x,y\in\mathbb{R}^n$, $\langle x,y\rangle:=x^\top y$ denotes the Euclidean inner product. For a matrix M, $\|M\|:=\sqrt{\lambda_{\max}(A^\top A)}$ is the induced 2-norm. Given the set $S\subseteq\mathbb{R}^n$, ∂S and $\mathrm{int}(S)$ represent the boundary and the interior of S, respectively.

2.2. PRELIMINARIES

We briefly recall some notions from hybrid dynamical systems that we will use to develop our results. Then, the problem statement is introduced along with some assumptions related to the optimization problem to be tackled in this chapter. We adapt the following definition of a hybrid control system from [36] that is sufficient in the context of this chapter.

Definition 2.2.1 (Hybrid control system)**.** *A time-invariant hybrid control system $\mathcal{H}$ comprises a controlled ODE and a jump (or a reset) rule introduced as:*

$$\left\{\begin{array}{lcll}\dot{x} &=& F\big(x,u(x)\big), & x\in\mathcal{C}\\ x^+ &=& G(x), & \textit{otherwise,}\end{array}\right. \qquad (\mathcal{H})$$

where x^+ is the state of the hybrid system after a jump, the function $u:\mathbb{R}^n\to\mathbb{R}^m$ denotes a feedback signal, the function $F:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ is the flow map, the set $\mathcal{C}\subseteq\mathbb{R}^n$ is the flow set, and the function $G:\partial\mathcal{C}\to\mathrm{int}(\mathcal{C})$ represents the jump map.

In hybrid dynamical systems, the notion of *Zeno behavior* refers to the phenomenon that an infinite number of jumps occur in a bounded time interval. We then call a solution trajectory of a hybrid dynamical system Zeno-free if the number of jumps within any finite time interval is bounded. The existence of a lower bound on the time interval between two consecutive jumps suffices to guarantee the Zeno-freeness of a solution

trajectory of a hybrid control system. Nonetheless, there exist solution concepts in the literature that accept Zeno behaviors, see for example [36–39] and the references therein.

Consider the following class of unconstrained optimization problems:

$$f^* := \min_{X \in \mathbb{R}^n} f(X), \tag{2.2}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is an objective function. We now formally state the main problem to be addressed in this chapter:

Problem 2.2.2. *Consider the unconstrained optimization problem (2.2) where the objective function f is twice differentiable. Given a positive scalar α, design a fast gradient-based method in the form of a hybrid control system ($\mathscr{H}$) with α-exponential convergence rate, i.e. for any initial condition $X(0)$ and any $t \geq 0$ we have*

$$f\big(X(t)\big) - f^* \leq e^{-\alpha t}\Big(f\big(X(0)\big) - f^*\Big),$$

where $\{X(t)\}_{t\geq 0}$ denotes the solution trajectory of the system ($\mathscr{H}$).

Assumption 2.2.3 (Regularity assumptions)**.** *We stipulate that the objective function $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable and fulfills the following:*

- *(Bounded Hessian) The Hessian of function f, denoted by $\nabla^2 f(x)$, is uniformly bounded, i.e.,*

$$-\ell_f I_n \preceq \nabla^2 f(x) \preceq L_f I_n, \tag{A1}$$

  where ℓ_f and L_f are non-negative constants.

- *(Gradient dominated) The function f satisfies the Polyak-Łojasiewicz inequality with a positive constant μ_f, i.e., for every x in $\mathbb{R}^n$ we have*

$$\frac{1}{2}\big\|\nabla f(x)\big\|^2 \geq \mu_f\big(f(x) - f^*\big), \tag{A2}$$

  where f^* is the minimum value of f on $\mathbb{R}^n$.

- *(Lipschitz Hessian) The Hessian of the function f is Lipschitz, i.e., for every x, y in $\mathbb{R}^n$ we have*

$$\big\|\nabla^2 f(x) - \nabla^2 f(y)\big\| \leq H_f \|x - y\|, \tag{A3}$$

  where H_f is a positive constant.

Remark 2.2.4 (Lipschitz gradient)**.** *Since the function f is twice differentiable, Assumption (A1) implies that the function ∇f is also Lipschitz with a positive constant L_f, i.e., for every x, y in $\mathbb{R}^n$ we have*

$$\big\|\nabla f(x) - \nabla f(y)\big\| \leq L_f \|x - y\|. \tag{2.3}$$

We now collect two remarks underlining some features of the set of functions that satisfy (A2).

Remark 2.2.5 (PL functions and invexity)**.** *The PL inequality in general does not imply the convexity of a function but rather the invexity of it. The notion of invexity was first introduced by [40]. The PL inequality (A2) implies that the suboptimality measure $f - f^*$ grows at most as a quadratic function of ∇f.*

Remark 2.2.6 (Non-uniqueness of stationary points)**.** *While the PL inequality does not require the uniqueness of the stationary points of a function (i.e., $\{x : \nabla f(x) = 0\}$), it ensures that all stationary points of the function f are global minimizers [41].*

We close our preliminary section with a couple of popular examples borrowed from [42].

Example 1 (PL functions). The composition of a strongly convex function and a linear function satisfies the PL inequality. This class includes a number of important problems such as least squares, i.e., $f(x) = \|Ax - b\|^2$ (obviously, strongly convex functions also satisfy the PL inequality). Any strictly convex function over a compact set satisfies the PL inequality. As such, the log-loss objective function in logistic regression, i.e., $f(x) = \Sigma_{i=1}^{n} \log\big(1 + \exp(b_i a_i^\top x)\big)$, locally satisfies the PL inequality.

2.3. MAIN RESULTS

In this section, the main results of this chapter are provided. We begin with introducing two types of structures for the hybrid system ($\mathscr{H}$) motivated by the dynamics of fast gradient methods [24]. Given a positive scalar α, these structures, indexed by **I** and **II**, enable achieving the rate of convergence $\mathcal{O}(e^{-\alpha t})$ in the suboptimality measure $f\big(X(t)\big) - f^*$. We then collect multiple remarks highlighting the shared implications of the two structures along with a naive type of time-discretization for these structures. The technical proofs are presented in Section 2.4. For notational simplicity, we introduce the notation $x := (x_1, x_2)$ such that the variables x_1 and x_2 represent the system trajectories X and $\dot{X}$, respectively.

2.3.1. STRUCTURE I: STATE-DEPENDENT DAMPING COEFFICIENT

The description of the first structure follows. We start with the flow map $F_{\mathbf{I}} : \mathbb{R}^{2n} \times \mathbb{R} \to \mathbb{R}^{2n}$ defined as

$$F_{\mathbf{I}}\big(x, u_{\mathbf{I}}(x)\big) = \begin{pmatrix} x_2 \\ -\nabla f(x_1) \end{pmatrix} + \begin{pmatrix} 0 \\ -x_2 \end{pmatrix} u_{\mathbf{I}}(x). \tag{2.4a}$$

Notice that $F_{\mathbf{I}}(\cdot,\cdot)$ is the state-space representation of a 2nd-order ODE. The feedback law $u_{\mathbf{I}} : \mathbb{R}^{2n} \to \mathbb{R}$ is given by

$$u_{\mathbf{I}}(x) = \alpha + \frac{\|\nabla f(x_1)\|^2 - \langle \nabla^2 f(x_1) x_2, x_2 \rangle}{\langle \nabla f(x_1), -x_2 \rangle}. \tag{2.4b}$$

Next, the candidate flow set $\mathscr{C}_{\mathbf{I}} \subset \mathbb{R}^{2n}$ is characterized by an admissible input interval $[\underline{u}_{\mathbf{I}}\ \overline{u}_{\mathbf{I}}]$, i.e.,

$$\mathscr{C}_{\mathbf{I}} = \big\{x \in \mathbb{R}^{2n} :\ u_{\mathbf{I}}(x) \in [\underline{u}_{\mathbf{I}}, , \overline{u}_{\mathbf{I}}]\big\}, \tag{2.4c}$$

where the interval bounds $\underline{u}_{\mathbf{I}}, \overline{u}_{\mathbf{I}}$ represent the range of admissible control values. Notice that the flow set $\mathscr{C}_{\mathbf{I}}$ is the domain in which the hybrid system $(\mathscr{H})$ can evolve continuously. Finally, we introduce the jump map $G_{\mathbf{I}} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ parameterized by a constant $\beta_{\mathbf{I}}$

$$G_{\mathbf{I}}(x) = \begin{pmatrix} x_1 \\ -\beta_{\mathbf{I}} \nabla f(x_1) \end{pmatrix}. \tag{2.4d}$$

The parameter $\beta_{\mathbf{I}}$ ensures that the range space of the jump map $G_{\mathbf{I}}$ is a strict subset of $\text{int}(\mathscr{C}_{\mathbf{I}})$. By construction, one can inspect that any neighborhood of the optimizer x_1^* has a non-empty intersection with the flow set $\mathscr{C}_{\mathbf{I}}$. That is, there always exist paths in the set $\mathscr{C}_{\mathbf{I}}$ that allow the continuous evolution of the hybrid system to approach arbitrarily close to the optimizer.

We are now in a position to formally present the main results related to the structure **I** given in (2.4). This theorem provides a framework to set the parameters $\underline{u}_{\mathbf{I}}$, $\overline{u}_{\mathbf{I}}$, and $\beta_{\mathbf{I}}$ in (2.4c) and (2.4d) in order to ensure the desired exponential convergence rate $\mathscr{O}(e^{-\alpha t})$.

Theorem 2.3.1 (Continuous-time convergence rate - **I**)**.** *Consider a positive scalar α and a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ satisfying Assumptions (A1) and (A2). Then, the solution trajectory of the hybrid control system $(\mathscr{H})$ with the respective parameters (2.4) starting from any initial condition $x_1(0)$ satisfies*

$$f\big(x_1(t)\big) - f^* \leq e^{-\alpha t}\Big(f\big(x_1(0)\big) - f^*\Big), \quad \forall\, t \geq 0, \tag{2.5}$$

if the scalars $\underline{u}_{\boldsymbol{I}}$, $\overline{u}_{\boldsymbol{I}}$, and $\beta_{\boldsymbol{I}}$ are chosen such that

$$\underline{u}_{\boldsymbol{I}} < \alpha + \beta_{\boldsymbol{I}}^{-1} - L_f \beta_{\boldsymbol{I}}, \tag{2.6a}$$
$$\overline{u}_{\boldsymbol{I}} > \alpha + \beta_{\boldsymbol{I}}^{-1} + \ell_f \beta_{\boldsymbol{I}}, \tag{2.6b}$$
$$\alpha \leq 2\mu_f \beta_{\boldsymbol{I}}. \tag{2.6c}$$

The next result establishes a key feature of the solution trajectories generated by the dynamics $(\mathscr{H})$ with the respective parameters (2.4), that the solution trajectories are indeed *Zeno*-free.

Theorem 2.3.2 (Zeno-free hybrid trajectories - **I**)**.** *Consider a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ satisfying Assumption 2.2.3, and the corresponding hybrid control system $(\mathscr{H})$ with the respective parameters (2.4) satisfying (2.6). Given the initial condition $\Big(x_1(0), -\beta_{\boldsymbol{I}} \nabla f\big(x_1(0)\big)\Big)$ the time between two consecutive jumps of the solution trajectory, denoted by $\tau_{\boldsymbol{I}}$, satisfies for any scalar $r > 1$*

$$\tau_{\boldsymbol{I}} \geq \log\left(\max\left\{\frac{a_1}{a_2 + a_3\Big\|\nabla f\big(x_1(0)\big)\Big\|} + 1, r^{1/\delta}\right\}\right), \tag{2.7}$$

where the constants involved are defined as

$$C := \frac{(\overline{u}_{\boldsymbol{I}} - \alpha) + \sqrt{(\overline{u}_{\boldsymbol{I}} - \alpha)^2 + 4L_f}}{2}, \tag{2.8a}$$

$$\delta := C + \max\{\overline{u}_{\boldsymbol{I}}, -\underline{u}_{\boldsymbol{I}}\}, \tag{2.8b}$$

$$\mathscr{L}_f := \max\{\ell_f, L_f\}, \tag{2.8c}$$

$$a_1 := \min\{\overline{u}_{\boldsymbol{I}} - (\alpha + \beta_{\boldsymbol{I}}^{-1} + \ell_f \beta_{\boldsymbol{I}}), (\alpha + \beta_{\boldsymbol{I}}^{-1} - L_f \beta_{\boldsymbol{I}}) - \underline{u}_{\boldsymbol{I}}\}, \tag{2.8d}$$

$$a_2 := rL_f \delta^{-1}(r\beta_{\boldsymbol{I}} C + 1) + \beta_{\boldsymbol{I}}^{-1} + (r^2 + r + 1)\beta_{\boldsymbol{I}} \mathscr{L}_f, \tag{2.8e}$$

$$a_3 := r^3 \beta_{\boldsymbol{I}}^2 H_f \delta^{-1}. \tag{2.8f}$$

Consequently, the solution trajectories are Zeno-free.

Remark 2.3.3 (Non-uniform inter-jumps - **I**)**.** *Notice that Theorem 2.3.2 suggests a lower-bound for the inter-jump interval $\tau_{\boldsymbol{I}}$ that depends on $\|\nabla f(x_1)\|$. In light of the fact that the solution trajectories converge to the optimal solutions, and as such $\nabla f(x_1)$ tends to zero, one can expect that the frequency at which the jumps occur reduces as the hybrid control system evolves in time.*

2.3.2. STRUCTURE II: STATE-DEPENDENT POTENTIAL COEFFICIENT

In this subsection, we first provide the structure **II** for the hybrid control system $(\mathscr{H})$. We skip the details of differences with the structure **I** and differ it to Subection 2.3.3 and Section 2.4. Consider the flow map $F_{\mathbf{II}} : \mathbb{R}^{2n} \times \mathbb{R} \to \mathbb{R}^{2n}$ given by

$$F_{\mathbf{II}}\big(x, u_{\mathbf{II}}(x)\big) = \begin{pmatrix} x_2 \\ -x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ -\nabla f(x_1) \end{pmatrix} u_{\mathbf{II}}(x), \tag{2.9a}$$

and the feedback law $u_{\mathbf{II}} : \mathbb{R}^{2n} \to \mathbb{R}$ given by

$$u_{\mathbf{II}}(x) = \frac{\langle \nabla^2 f(x_1) x_2, x_2 \rangle + (1-\alpha)\langle \nabla f(x_1), -x_2 \rangle}{\|\nabla f(x_1)\|^2}. \tag{2.9b}$$

The candidate flow set $\mathscr{C}_{\mathbf{II}} \subset \mathbb{R}^{2n}$ is parameterized by an admissible interval $[\underline{u}_{\mathbf{II}}\ \overline{u}_{\mathbf{II}}]$ as follows:

$$\mathscr{C}_{\mathbf{II}} = \left\{ x \in \mathbb{R}^{2n} : \ u_{\mathbf{II}}(x) \in [\underline{u}_{\mathbf{II}}, \overline{u}_{\mathbf{II}}] \right\}. \tag{2.9c}$$

Parameterized in a constant $\beta_{\mathbf{II}}$, the jump map $G_{\mathbf{II}} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is given by

$$G_{\mathbf{II}}(x) = \begin{pmatrix} x_1 \\ -\beta_{\mathbf{II}} \nabla f(x_1) \end{pmatrix}. \tag{2.9d}$$

Theorem 2.3.4 (Continuous-time convergence rate - **II**)**.** *Consider a positive scalar α and a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ satisfying Assumptions (A1) and (A2). Then, the solution trajectory of the hybrid control system $(\mathscr{H})$ with the respective parameters* (2.9) *starting*

from any initial condition $x_1(0)$ *satisfies the inequality* (2.5) *if the scalars* $\underline{u}_{\boldsymbol{II}}$, $\overline{u}_{\boldsymbol{II}}$, *and* $\beta_{\boldsymbol{II}}$ *are chosen such that*

$$\underline{u}_{\boldsymbol{II}} < -\ell_f \beta_{\boldsymbol{II}}^2 + (1-\alpha)\beta_{\boldsymbol{II}}, \tag{2.10a}$$

$$\overline{u}_{\boldsymbol{II}} > L_f \beta_{\boldsymbol{II}}^2 + (1-\alpha)\beta_{\boldsymbol{II}}, \tag{2.10b}$$

$$\alpha \leq 2\mu_f \beta_{\boldsymbol{II}}. \tag{2.10c}$$

Theorem 2.3.5 (Zeno-free hybrid trajectories - **II**)**.** *Consider a smooth function* $f : \mathbb{R}^n \to \mathbb{R}$ *satisfying Assumptions* (A1) *and* (A2)*, and the hybrid control system* $(\mathscr{H})$ *with the respective parameters* (2.9) *satisfying* (2.10)*. Given the initial condition* $\Big(x_1(0), -\beta_{\boldsymbol{II}} \nabla f\big(x_1(0)\big)\Big)$ *the time between two consecutive jumps of the solution trajectory, denoted by* $\tau_{\boldsymbol{II}}$*, satisfies for any scalar* $r \in (0,1)$

$$\tau_{\boldsymbol{II}} \geq \min\left\{r\omega^{-1}, \delta(b_1+b_2)^{-1}\right\}, \tag{2.11}$$

where the involved scalars are defined as

$$\begin{aligned}
\delta &:= \min\left\{\overline{u}_{\boldsymbol{II}} - (L_f\beta_{\boldsymbol{II}}^2 + (1-\alpha)\beta_{\boldsymbol{II}}), (-\ell_f\beta_{\boldsymbol{II}}^2 + (1-\alpha)\beta_{\boldsymbol{II}}) - \underline{u}_{\boldsymbol{II}}\right\},\\
U &:= \max\{\overline{u}_{\boldsymbol{II}}, -\underline{u}_{\boldsymbol{II}}\},\\
\mathscr{L}_f &:= \max\{\ell_f, L_f\},\\
\omega &:= \mathscr{L}_f(\beta_{\boldsymbol{II}}^2 + \beta_{\boldsymbol{II}}U)^{\frac{1}{2}},\\
b1 &:= \frac{2\mathscr{L}_f\beta_{\boldsymbol{II}}\big(U + \omega(\beta_{\boldsymbol{II}}+U)\big)}{(1-r)^3},\\
b_2 &:= |\alpha-1|\frac{2\omega\beta_{\boldsymbol{II}}}{(1-r)^3} + |\alpha-1|\alpha\beta_{\boldsymbol{II}}(1+r).
\end{aligned}$$

Thus, the solution trajectories are Zeno-free.

Remark 2.3.6 (Uniform inter-jumps - **II**)**.** *Notice that unlike Theorem 2.3.2, the derived lower-bound for the inter-jump interval* $\tau_{\boldsymbol{II}}$ *is uniform in the sense that the bound is independent of* $\|\nabla f(x_1)\|$*. Furthermore, the regularity requirement on the function* f *is weaker than the one used in Theorem 2.3.2, i.e., the function* f *is not required to satisfy the Assumption* (A3)*.*

Notice that the main differences between the structures (2.4), (2.9) lie in the flow maps and the feedback laws. On the other hand, these structures share the key feature of enabling an α-exponential convergence rate for the hybrid system $(\mathscr{H})$ through their corresponding control inputs. The reason explaining the aforementioned points is deferred until later in Section 2.4.

2.3.3. FURTHER DISCUSSIONS

In what follows, we collect several remarks regarding the common features of the proposed structures. Then, we apply the *forward-Euler* method of time-discretization to these structures of the hybrid control system $(\mathscr{H})$. The proposed discretizations guarantee an exponential rate of convergence in the suboptimality measure $f(x_1^k) - f^*$, where k is the iteration index.

Remark 2.3.7 (Weaker regularity than strong convexity)**.** *The PL inequality is a weaker requirement than the strong convexity, which is often assumed in similar contexts [24, 26, 27]. It is worth noting that such a condition has also been used in the context of 1st-order algorithms [42].*

Remark 2.3.8 (Hybrid embedding of restarting)**.** *The hybrid frameworks intrinsically capture restarting schemes through the jump map. The schemes are a weighted gradient where the weight factor $\beta_{\mathbf{I}}$ or $\beta_{\mathbf{II}}$ is essentially characterized by the given data α, μ_f, ℓ_f, and L_f. One may inspect that the constant $\beta_{\mathbf{I}}$ or $\beta_{\mathbf{II}}$ can be in fact introduced as a state-dependent weight factor to potentially improve the performance. Nonetheless, for the sake of simplicity of exposition, we do not pursue this level of generality in this chapter.*

Remark 2.3.9 (2nd-order information)**.** *Although our proposed frameworks require 2nd-order information, i.e., the Hessian $\nabla^2 f$, this requirement only appears in a mild form as an evaluation in the same spirit as the modified Newton step proposed in [43]. Furthermore, we emphasize that our results still hold true if one replaces $\nabla^2 f(x_1)$ with its upper-bound $L_f I_n$ following essentially the same analysis. For further details we refer the reader to the proof of Theorem 2.3.4.*

Remark 2.3.10 (Fundamental limits on control input)**.** *An implication of Theorem 2.3.4 is that if the desired convergence rate $\alpha > \left(\frac{2\mu_f}{2\mu_f+\ell_f}\right)$, it is then required to choose $\underline{u}_{\mathbf{II}} < 0$, indicating that the system may need to receive energy through a negative damping. On a similar note, Theorem 2.3.1 asserts that the upper bound requires $\overline{u}_{\mathbf{I}} > \alpha$, and if $\alpha > \left(\frac{2\mu_f}{\sqrt{\max\{L_f-2\mu_f,0\}}}\right)$, we then have to set $\underline{u}_{\mathbf{I}} < 0$.*

2.3.4. DISCRETE-TIME DYNAMICS

In the next result, we show that if one applies the forward-Euler method on the two proposed structures properly, the resulting discrete-time hybrid control systems possess exponential convergence rates. Suppose $i \in \{\mathbf{I}, \mathbf{II}\}$ and let us denote by s the time-discretization step size. Consider the discrete-time hybrid control system

$$\mathcal{H}_{d,i} := \begin{cases} x^{k+1} = F_{d,i}\big(x^k, u_{d,i}(x^k)\big), & x^k \in \mathcal{C}_{d,i} \\ x^{k+1} = G_{d,i}(x^k), & \text{otherwise,} \end{cases} \tag{2.12}$$

where $F_{d,i}$, $G_{d,i}$, and $\mathcal{C}_{d,i}$ are the flow map, the jump map, and the flow set, respectively. The discrete flow map $F_{d,i} : \mathbb{R}^{2n} \times \mathbb{R} \to \mathbb{R}^{2n}$ is given by

$$F_{d,i}\big(x^k, u_{d,i}(x^k)\big) = x^k + sF_i\big(x^k, u_i(x^k)\big),\ i \in \{\mathbf{I}, \mathbf{II}\}, \tag{2.13a}$$

where F_i and u_i are defined in (2.4a) and (2.4b), or (2.9a) and (2.9b) based on the considered structure i. The discrete flow set $\mathcal{C}_{d,i} \subset \mathbb{R}^{2n}$ is defined as

$$\mathcal{C}_{d,i} := \big\{(x_1^k, x_2^k) \in \mathbb{R}^{2n} : c_1\|x_2^k\|^2 \le \|\nabla f(x_1^k)\|^2 \le c_2\langle \nabla f(x_1^k), -x_2^k\rangle\big\}, \tag{2.13b}$$

and, c_1 and c_2 are two positive scalars. The discrete jump map $G_{d,i} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is given by $G_{d,i}(x^k) = \big((x^k)^\top, -\beta\nabla^\top f(x^k)\big)^\top$.

It is evident in the flow sets $\mathscr{C}_{d,i}$ of the discrete-time dynamics that these sets are no longer defined based on admissible input intervals. The reason has to do with the difficulties that arise from appropriately discretizing the control inputs $u_{\mathbf{I}}$ and $u_{\mathbf{II}}$. Nonetheless, the next result guarantees exponential rate of convergence of the discrete-time control system (2.12) with either of the respective structure **I** or **II**, by introducing a mechanism to set the scalars c_1, c_2, and β.

Theorem 2.3.11 (Stable discretization - **I & II**)**.** *Consider a smooth function $f:\mathbb{R}^n\to\mathbb{R}$ satisfying Assumptions* (A1) *and* (A2)*. The solution trajectory of the discrete-time hybrid control system (2.12) with the respective structure $i\in\{\mathbf{I},\mathbf{II}\}$ and starting from any initial condition x_1^0, satisfies*

$$f(x_1^{k+1})-f^*\leq\lambda(s,c_1,c_2,\beta)\big(f(x_1^k)-f^*\big), \tag{2.14}$$

with $\lambda(s,c_1,c_2,\beta)\in(0,1)$ given by

$$\lambda(s,c_1,c_2,\beta):=1+2\mu_f\big(-\frac{s}{c_2}+\frac{L_f}{2c_1}s^2\big), \tag{2.15}$$

if the parameters s, c_1 ,c_2, and β satisfy

$$\sqrt{c_1}\leq c_2, \tag{2.16a}$$
$$\beta^2 c_1\leq 1\leq\beta c_2, \tag{2.16b}$$
$$c_2L_fs<2c_1. \tag{2.16c}$$

Remark 2.3.12 (Naive discretization)**.** *We would like to emphasize that the exponential convergence of the proposed discretization method solely depends on the dynamics x_1 and the properties of the objective function f. Thus, we deliberately avoid labeling the scalars c_1, c_2, and β by the structure index i. Crucially, the structures of the control laws do not impact the relations* (2.16) *in Theorem 2.3.11, see Subsection 2.4.5 for more details. In light of the above facts, we believe that a more in-depth analysis of the dynamics along with the control structures may provide a more intelligent way to improve the discretization result of Theorem 2.3.11.*

Corollary 2.3.13 (Optimal guaranteed rate)**.** *The optimal convergence rate guaranteed by Theorem 2.3.11 for the discrete-time dynamics is $\lambda^*:=\big(1-\frac{\mu_f}{L_f}\big)$ and*

$$\sqrt{c_1^*}=c_2^*=\frac{1}{\beta^*}=L_fs^*.$$

The pseudocode to implement the above corollary is presented in Algorithm 1 using the discrete-time dynamics (2.12) with the respective parameters **I** or **II**.

2.4. TECHNICAL PROOFS

2.4.1. PROOF OF THEOREM 2.3.1

We start with an explanation on why the chosen structure for $u_{\mathbf{I}}(x)$ guarantees the desired convergence rate α. Let us define the set

$$\mathscr{E}_\alpha:=\Big\{x\in\mathbb{R}^{2n}:\alpha\big(f(x_1)-f^*\big)<\langle\nabla f(x_1),-x_2\rangle\Big\}.$$

Algorithm 1 State-dependent fast gradient method

Input: data x_1^0, ℓ_f, L_f, μ_f, $\alpha \in \mathbb{R}^+$, $k_{\max} \in \mathbb{N}^+$, $i \in \{\mathbf{I}, \mathbf{II}\}$
Set: $\sqrt{c_1} = c_2 = \beta^{-1} = L_f s$, $x_2^0 = -\beta \nabla f(x_1^0)$
    $x^0 = (x_1^0, x_2^0)$
for $k = 1$ **to** $k_{\max}$ **do**
    if $c_1 \|x_2^k\|^2 \leq \|\nabla f(x_1^k)\|^2 \leq c_2 \langle \nabla f(x_1^k), -x_2^k \rangle$ **then**
        $x^{k+1} \leftarrow F_{d,i}(x^k)$
    else
        $x^{k+1} \leftarrow G_{d,i}(x^k)$
    end if
end for

In the first step, we argue that the objective function f decreases at the rate α (i.e., (2.5)) along any solution trajectory of the dynamical system (2.4a) that is contained in the set $\mathscr{E}_\alpha$. To see this, observe that if $\big(x_1(t), x_2(t)\big) \in \mathscr{E}_\alpha$, we then have

$$\frac{d}{dt}\Big(f\big(x_1(t)\big) - f^*\Big) = \big\langle \nabla f\big(x_1(t)\big), x_2(t)\big\rangle \leq -\alpha\big(f(x_1) - f^*\big).$$

The direct application of Gronwall's inequality, see [44, Lemma A.1], to the above inequality yields the desired convergence claim (2.5). In light of the above observation, it suffices to ensure that the solution trajectory does not leave the set $\mathscr{E}_\alpha$. Let us define the quantity

$$\sigma(t) := \langle \nabla f\big(x_1(t)\big), x_2(t)\rangle + \alpha\Big(f\big(x_1(t)\big) - f^*\Big).$$

By definition, if $\sigma(t) < 0$, it is then readily guaranteed that $\big(x_1(t), x_2(t)\big) \in \mathscr{E}_\alpha$. By virtue of this implication, if $\dot\sigma(t) \leq 0$ along the solution trajectory of (2.4a), we ensure that the value of $\sigma(t)$ does not increase, and as such

$$\big(x_1(t), x_2(t)\big) \in \mathscr{E}_\alpha,\ \forall t \geq 0 \iff \big(x_1(0), x_2(0)\big) \in \mathscr{E}_\alpha.$$

To ensure non-positivity property of $\dot\sigma(t)$, note that we have

$$\begin{aligned}\dot\sigma(t) &= \langle \nabla^2 f\big(x_1(t)\big)x_2, x_2(t)\rangle + \langle \nabla f\big(x_1(t)\big), \dot x_2(t)\rangle + \alpha\langle \nabla f\big(x_1(t)\big), x_2(t)\rangle \\ &= \langle \nabla^2 f\big(x_1(t)\big)x_2(t), x_2(t)\rangle - \|\nabla f\big(x_1(t)\big)\|^2 + \Big(\alpha - u_{\mathbf{I}}\big(x(t)\big)\Big)\langle \nabla f\big(x_1(t)\big), x_2(t)\rangle = 0,\end{aligned}$$

where the last equality follows from the definition of the proposed control law (2.4b). It is worth noting that one can simply replace the information of the Hessian $\nabla^2 f\big(x_1(t)\big)$ with the upper bound L_f and still arrive at the desired inequality, see also Remark 2.3.9 with regards to the 1st-order information oracle. Thus far, we have shown how the designed feedback control preserves the α-rate of convergence along the continuous flow of the hybrid system. Consider the initial state $x_2(0) = -\beta \nabla f\big(x_1(0)\big)$. To ensure $x(0) \in \mathscr{E}_\alpha$, notice

that

$$\begin{aligned}\alpha\Big(f\big(x_1(0)\big)-f^*\Big) &\le \frac{\alpha}{2\mu_f}\big\|\nabla f\big(x_1(0)\big)\big\|^2\\ &= \frac{\alpha}{2\mu_f\beta}\langle -x_2(0),\nabla f\big(x_1(0)\big)\rangle\\ &\le \langle \nabla f\big(x_1(0)\big),-x_2(0)\rangle,\end{aligned}$$

where in the first line we use (A2), and in the last line the condition (2.6c) is employed. Introducing the proposed $x_2(0)$ as the jump x^+ one can see that the range space of the jump map (2.4d) is indeed contained in the set $\mathscr{E}_\alpha$. Finally, we need to ensure that such a jump policy is well-defined, that is the trajectory lands in the interior of the flow set $\mathscr{C}_{\mathbf{I}}$ defined as in (2.4c), i.e., the control values also belong to the admissible set $[\underline{u}_{\mathbf{I}},\overline{u}_{\mathbf{I}}]$. In this view, we only need to take the initial control value into consideration, as the switching law is continuous in the states and serves the purpose by design. Suppose that $x\in\mathscr{C}_{\mathbf{I}}$, we then have the sufficient requirements

$$\underline{u}_{\mathbf{I}}<\alpha+\frac{\|\nabla f(x_1^+)\|^2-L_f\beta^2\|\nabla f(x_1^+)\|^2}{\beta\|\nabla f(x_1^+)\|^2}\le u_{\mathbf{I}}(x^+)\le\alpha+\frac{\|\nabla f(x_1^+)\|^2+\ell_f\beta^2\|\nabla f(x_1^+)\|^2}{\beta\|\nabla f(x_1^+)\|^2}<\overline{u}_{\mathbf{I}},$$

where the relations (2.4b) and (A1) are considered. Canceling the term $\|\nabla f(x_1^+)\|^2$ concludes the sufficient requirements in (2.6a) and (2.6b).

2.4.2. PROOF OF THEOREM 2.3.2

In this subsection, we first set the stage by providing two intermediate results regarding the properties of dynamics of the hybrid control system ($\mathscr{H}$) with the respective parameters (2.4). We then employ these facts to formally state the proof of Theorem 2.3.2. The next lemma reveals a relation between $\nabla f(x_1)$ and x_2 along the trajectories of the hybrid control system. In this subsection, for the sake of brevity we denote $x_1(t)$ and $x_1(0)$ by x_1 and $x_{1,0}$, respectively. We adapt the same change of notation for x_2 and x, as well.

Lemma 2.4.1 (Velocity lower bound)**.** *Consider the continuous-time hybrid control system ($\mathscr{H}$) with the respective parameters (2.4) satisfying (2.6) where the function f satisfies Assumptions (A1) and (A2). Then, we have*

$$\big\|\nabla f(x_1)\big\|\le C\|x_2\|, \tag{2.17}$$

where C is given by (2.8a).

Proof. Notice that, by the definition of the control law and the upper bound condition $u_{\mathbf{I}}(x)\le\overline{u}_{\mathbf{I}}$, we have

$$\big\|\nabla f(x_1)\big\|^2-\langle\nabla^2 f(x_1)x_2,x_2\rangle\le(\overline{u}_{\mathbf{I}}-\alpha)\langle\nabla f(x_1),-x_2\rangle\le(\overline{u}_{\mathbf{I}}-\alpha)\big\|\nabla f(x_1)\big\|\cdot\|x_2\|,$$

where the second inequality follows from the Cauchy-Schwarz inequality. Since the function f satisfies Assumption (A1), one can infer that

$$\big\|\nabla f(x_1)\big\|^2-L_f\|x_2\|^2\le(\overline{u}_{\mathbf{I}}-\alpha)\big\|\nabla f(x_1)\big\|\cdot\|x_2\|,$$

which in turn can be reformulated into

$$\frac{\|\nabla f(x_1)\|^2}{\|x_2\|^2} - (\overline{u}_{\mathbf{I}} - \alpha)\frac{\|\nabla f(x_1)\|}{\|x_2\|} - L_f \leq 0. \tag{2.18}$$

2

Defining the variable $y := \|\nabla f(x_1)\|/\|x_2\|$, the inequality (2.18) becomes the quadratic inequality $y^2 - (\overline{u}_{\mathbf{I}} - \alpha)y - L_f \leq 0$. Taking into account that $y \geq 0$, it then follows from (2.17) that

$$y = \frac{\|\nabla f(x_1)\|}{\|x_2\|} \leq \frac{(\overline{u}_{\mathbf{I}} - \alpha) + \sqrt{(\overline{u}_{\mathbf{I}} - \alpha)^2 + 4L_f}}{2} =: C.$$

This concludes the proof of Lemma 2.4.1. □

In the following, we provide a result that indicates the variation of norms x_1 and x_2, along the trajectories of the hybrid control system, are bounded in terms of time while they evolve according to the continuous mode. Since the hybrid control system is time-invariant, such bounds can be generalized to all inter-jump intervals.

Lemma 2.4.2 (Trajectory growth rate)**.** *Suppose that the same conditions as specified in Lemma 2.4.1 hold, and the hybrid control system (ℋ), (2.4) starts from the initial condition* $\big(x_{1,0}, -\beta_{\boldsymbol{I}}\nabla f(x_{1,0})\big)$ *for some* $x_{1,0} \in \mathbb{R}^n$*. Then*

$$\|x_1 - x_{1,0}\| \leq \delta^{-1}\|x_{2,0}\|\big(e^{\delta t} - 1\big), \tag{2.19a}$$

$$\|x_2 - x_{2,0}\| \leq \|x_{2,0}\|\big(e^{\delta t} - 1\big), \tag{2.19b}$$

where δ *is given by (2.8b).*

Proof. Using the flow dynamics (2.4a) we obtain

$$\frac{d}{dt}\|x_2\| \leq \left\|\frac{d}{dt}x_2\right\| \leq \|\nabla f(x_1)\| + \big|u_{\mathbf{I}}(x)\big| \cdot \|x_2\| \leq (C + \max\{\overline{u}_{\mathbf{I}}, -\underline{u}_{\mathbf{I}}\})\|x_2\| = \delta\|x_2\|. \tag{2.20}$$

The inequality (2.20) implies that

$$\|x_2\| \leq \|x_{2,0}\|e^{\delta t}. \tag{2.21}$$

Furthermore, notice that

$$\frac{d}{dt}\|x_1 - x_{1,0}\| \leq \left\|\frac{d}{dt}(x_1 - x_{1,0})\right\| = \|x_2\|.$$

Integrating the two sides of the above inequality leads to

$$\|x_1 - x_{1,0}\| \leq \int_0^t \|x_2(s)\|\, ds \leq \int_0^t \|x_{2,0}\|e^{\delta s}\, ds = \frac{\|x_{2,0}\|}{\delta}\big(e^{\delta t} - 1\big),$$

in which we made use of (2.21). Hence, the inequality (2.19a) in Lemma 2.4.1 is concluded. Next, we shall establish the inequality (2.19b). Note that

$$\begin{aligned}\frac{d}{dt}\|x_2 - x_{2,0}\| \leq \left\|\frac{d}{dt}(x_2 - x_{2,0})\right\| = \left\|\frac{d}{dt}x_2\right\| &\leq \delta\|x_2\| \\ &\leq \delta\|x_2 - x_{2,0}\| + \delta\|x_{2,0}\|.\end{aligned}$$

Applying Grownwall's inequality [44, Lemma A.1] then leads to the inequality (2.19b). The claims in Lemma 2.4.2 follow. □

Proof of Theorem 2.3.2: The proof comprises five steps, and the key part is to guarantee that during the first inter-jump interval the quantity $\left|u_{\mathbf{I}}(x) - u_{\mathbf{I}}(x_{,0})\right|$ is bounded by a continuous function $\phi\Big(t, \big\|\nabla f(x_{1,0})\big\|\Big)$, which is exponential in its first argument and linear in its second argument. Then, it follows from the continuity of the function ϕ that the solution trajectories of the hybrid control system are Zeno-free.

Step 1: Let us define $g(t) := \langle \nabla f(x_1), -x_2\rangle$. We now compute the derivative of $g(t)$ along the trajectories of the hybrid control system $(\mathscr{H})$, (2.4) during the first inter-jump interval, i.e.,

$$\begin{aligned}\frac{d}{dt}g(t) &= \langle \nabla^2 f(x_1)x_2, -x_2\rangle + \langle \nabla f(x_1), u_{\mathbf{I}}(x)x_2 + \nabla f(x_1)\rangle \\ &= -\langle \nabla^2 f(x_1)x_2, x_2\rangle + \big\|\nabla f(x_1)\big\|^2 + u_{\mathbf{I}}(x)\langle \nabla f(x_1), x_2\rangle \\ &= -\alpha\langle \nabla f(x_1), -x_2\rangle = -\alpha\ g(t).\end{aligned}$$

According to the above discussion and considering the initial state $x_{2,0} = -\beta_{\mathbf{I}}\nabla f(x_{1,0})$, it follows that

$$\langle \nabla f(x_1), -x_2\rangle = \beta_{\mathbf{I}}\big\|\nabla f(x_{1,0})\big\|^2 e^{-\alpha t}. \tag{2.22}$$

Step 2: The quantity $\left|e^{\alpha t}\big\|\nabla f(x_1)\big\|^2 - \big\|\nabla f(x_{1,0})\big\|^2\right|$ is bounded along the trajectories of the hybrid control system $(\mathscr{H})$ with the respective parameters (2.4) during the first inter-jump interval, i.e.,

$$\begin{aligned}\left|e^{\alpha t}\big\|\nabla f(x_1)\big\|^2 - \big\|\nabla f(x_{1,0})\big\|^2\right| &= \left|e^{\alpha t}\big\|\nabla f(x_1)\big\|^2 - (e^{\alpha t} - e^{\alpha t} + 1)\big\|\nabla f(x_{1,0})\big\|^2\right| \\ &\overset{\text{(i)}}{\leq} e^{\alpha t}\left|\big\|\nabla f(x_1)\big\|^2 - \big\|\nabla f(x_{1,0})\big\|^2\right| + (e^{\alpha t} - 1)\big\|\nabla f(x_{1,0})\big\|^2 \\ &= e^{\alpha t}\left|\big\langle \nabla f(x_1) - \nabla f(x_{1,0}), \nabla f(x_1) + \nabla f(x_{1,0})\big\rangle\right| \\ &\quad + (e^{\alpha t} - 1)\big\|\nabla f(x_{1,0})\big\|^2 \\ &\overset{\text{(ii)}}{\leq} e^{\alpha t}\big\|\nabla f(x_1) - \nabla f(x_{1,0})\big\| \cdot \big\|\nabla f(x_1) + \nabla f(x_{1,0})\big\| \\ &\quad + (e^{\alpha t} - 1)\big\|\nabla f(x_{1,0})\big\|^2 \\ &\overset{\text{(iii)}}{\leq} e^{\alpha t}L_f\|x_1 - x_{1,0}\| \cdot \big(\beta_{\mathbf{I}}Ce^{\delta t} + 1\big)\frac{\|x_{2,0}\|}{\beta_{\mathbf{I}}} \\ &\quad + \big(e^{\alpha t} - 1\big)\frac{\|x_{2,0}\|^2}{\beta_{\mathbf{I}}^2} \\ &\overset{\text{(iv)}}{\leq} e^{\alpha t}L_f\big(e^{\delta t} - 1\big)\frac{\|x_{2,0}\|}{\delta} \cdot \big(\beta_{\mathbf{I}}Ce^{\delta t} + 1\big)\frac{\|x_{2,0}\|}{\beta_{\mathbf{I}}} \\ &\quad + \big(e^{\alpha t} - 1\big)\frac{\|x_{2,0}\|^2}{\beta_{\mathbf{I}}^2} \\ &= \left(\frac{L_f}{\delta\beta_{\mathbf{I}}}e^{\alpha t}\big(\beta_{\mathbf{I}}Ce^{\delta t} + 1\big)\big(e^{\delta t} - 1\big) + \frac{1}{\beta_{\mathbf{I}}^2}\big(e^{\alpha t} - 1\big)\right)\|x_{2,0}\|^2,\end{aligned}$$

where we made use of the triangle inequality in the inequality (i), the Cauchy-Schwarz inequality in the inequality (ii), Assumption (A1) and its consequence in Remark 2.2.4 along with the triangle inequality in the inequality (iii), and the inequality (2.19a) in the inequality (iv), respectively.

Step 3: Observe that

$$
\begin{aligned}
&\left|e^{\alpha t}\langle\nabla^2 f(x_1)x_2,x_2\rangle-\langle\nabla^2 f(x_{1,0})x_{2,0},x_{2,0}\rangle\right|\\
&\quad=\Big|e^{\alpha t}\big\langle\big[\nabla^2 f(x_1)-\nabla^2 f(x_{1,0})+\nabla^2 f(x_{1,0})\big]x_2,x_2\big\rangle\\
&\qquad-\big(e^{\alpha t}-e^{\alpha t}+1\big)\langle\nabla^2 f(x_{1,0})x_{2,0},x_{2,0}\rangle\Big|\\
&\quad=\Big|e^{\alpha t}\big\langle\big[\nabla^2 f(x_1)-\nabla^2 f(x_{1,0})\big]x_2,x_2\big\rangle\\
&\qquad+e^{\alpha t}\langle\nabla^2 f(x_{1,0})x_2,x_2\rangle-e^{\alpha t}\langle\nabla^2 f(x_{1,0})x_2,x_2\rangle\\
&\qquad+\big(e^{\alpha t}-1\big)\langle\nabla^2 f(x_{1,0})x_{2,0},x_{2,0}\rangle\Big|\\
&\quad\overset{\text{(i)}}{\leq}e^{\alpha t}\Big|\big\langle\big[\nabla^2 f(x_1)-\nabla^2 f(x_{1,0})\big]x_2,x_2\big\rangle\Big|\\
&\qquad+e^{\alpha t}\Big|\langle\nabla^2 f(x_{1,0})x_2,x_2\rangle-\langle\nabla^2 f(x_{1,0})x_{2,0},x_{2,0}\rangle\Big|\\
&\qquad+\big(e^{\alpha t}-1\big)\Big|\langle\nabla^2 f(x_{1,0})x_{2,0},x_{2,0}\rangle\Big|\\
&\quad\overset{\text{(ii)}}{\leq}e^{\alpha t}H_f\|x_1-x_{1,0}\|\cdot\|x_2\|^2\\
&\qquad+e^{\alpha t}\Big|\big\langle\nabla^2 f(x_{1,0})\big[x_2-x_{2,0}\big],x_2+x_{2,0}\big\rangle\Big|\\
&\qquad+\mathscr{L}_f\|x_{2,0}\|^2\big(e^{\alpha t}-1\big),
\end{aligned}
$$

where the inequality (i) follows from the triangle inequality, and the inequality (ii) is an immediate consequence of Assumptions (A3) and (A1), recalling $\mathscr{L}_f=\max\{\ell_f,L_f\}$. According to the above analysis, one can deduce that

$$
\begin{aligned}
&\left|e^{\alpha t}\langle\nabla^2 f(x_1)x_2,x_2\rangle-\langle\nabla^2 f(x_{1,0})x_{2,0},x_{2,0}\rangle\right|\\
&\overset{\text{(i)}}{\leq}e^{\alpha t}H_f\frac{\|x_{2,0}\|}{\delta}\big(e^{\delta t}-1\big)\cdot e^{2\delta t}\|x_{2,0}\|^2\\
&\quad+e^{\alpha t}\mathscr{L}_f\|x_2-x_{2,0}\|\cdot\|x_2+x_{2,0}\|+\big(e^{\alpha t}-1\big)\mathscr{L}_f\|x_{2,0}\|^2\\
&\overset{\text{(ii)}}{\leq}\frac{H_f}{\delta}e^{(\alpha+2\delta)t}\big\|x_2(0)\big\|^3\cdot\big(e^{\delta t}-1\big)\\
&\quad+e^{\alpha t}\mathscr{L}_f\big(e^{\delta t}-1\big)\|x_{2,0}\|\cdot\big(e^{\delta t}+1\big)\|x_{2,0}\|\\
&\quad+\mathscr{L}_f\|x_{2,0}\|^2\big(e^{\alpha t}-1\big)\\
&=\Big(\big(H_f/\delta\big)\,e^{(\alpha+2\delta)t}\|x_{2,0}\|\cdot\big(e^{\delta t}-1\big)\\
&\quad+\mathscr{L}_f\big(e^{(\alpha+\delta)t}+e^{\alpha t}\big)\big(e^{\delta t}-1\big)+\mathscr{L}_f\big(e^{\alpha t}-1\big)\Big)\|x_{2,0}\|^2,
\end{aligned}
$$

where we made use of the inequality (2.19a), the inequality (2.19b), and the triangle inequality in the inequality (i), and the inequality (2.19b) and the triangle inequality in the inequality (ii), respectively.

Step 4: We now study the input variation $\left|u_{\mathbf{I}}(x)-u_{\mathbf{I}}(x_{,0})\right|$ along the solution trajectories of the hybrid control system $(\mathscr{H})$, (2.4) during the first inter-jump interval. Observe that

$$
\begin{aligned}
&\left|u_{\mathbf{I}}(x)-u_{\mathbf{I}}(x_{,0})\right| \\
&= \Big| \frac{\big\|\nabla f(x_1)\big\|^2 - \langle \nabla^2 f(x_1)x_2(t), x_2\rangle}{\langle \nabla f(x_1), -x_2\rangle} - \frac{\big\|\nabla f(x_{1,0})\big\|^2 - \langle \nabla^2 f(x_{1,0})x_{2,0}, x_{2,0}\rangle}{\langle \nabla f(x_{1,0}), -x_{2,0}\rangle}\Big| \\
&= \Big| \frac{\big\|\nabla f(x_1)\big\|^2}{\beta_{\mathbf{I}}\big\|\nabla f(x_{1,0})\big\|^2 e^{-\alpha t}} - \frac{\langle \nabla^2 f(x_1)x_2, x_2\rangle}{\beta_{\mathbf{I}}\big\|\nabla f(x_{1,0})\big\|^2 e^{-\alpha t}} - \frac{\big\|\nabla f(x_{1,0})\big\|^2}{\beta_{\mathbf{I}}\big\|\nabla f(x_{1,0})\big\|^2} + \frac{\langle \nabla^2 f(x_{1,0})x_{2,0}, x_{2,0}\rangle}{\beta_{\mathbf{I}}\big\|\nabla f(x_{1,0})\big\|^2}\Big| \\
&\overset{\text{(i)}}{\leq} \frac{1}{\beta_{\mathbf{I}}\big\|\nabla f(x_{1,0})\big\|^2}\Big|e^{\alpha t}\big\|\nabla f(x_1)\big\|^2 - \big\|\nabla f(x_{1,0})\big\|^2\Big| \\
&\quad + \frac{1}{\beta_{\mathbf{I}}\big\|\nabla f(x_{1,0})\big\|^2}\Big|e^{\alpha t}\langle \nabla^2 f(x_1)x_2, x_2\rangle - \langle \nabla^2 f(x_{1,0})x_{2,0}, x_{2,0}\rangle\Big| \\
&\overset{\text{(ii)}}{=} \frac{\beta_{\mathbf{I}}}{\|x_{2,0}\|^2}\Big|e^{\alpha t}\big\|\nabla f(x_1)\big\|^2 - \big\|\nabla f(x_{1,0})\big\|^2\Big| \\
&\quad + \frac{\beta_{\mathbf{I}}}{\|x_{2,0}\|^2}\Big|e^{\alpha t}\langle \nabla^2 f(x_1)x_2, x_2\rangle - \langle \nabla^2 f(x_{1,0})x_{2,0}, x_{2,0}\rangle\Big|,
\end{aligned}
$$

where we made use of the triangle inequality in the inequality (i) and the relation (2.22) in the equality (ii), respectively. Based on the above discussion, we then conclude that

$$
\begin{aligned}
\left|u_{\mathbf{I}}(x)-u_{\mathbf{I}}(x_{,0})\right| &\overset{\text{(i)}}{\leq} \frac{\beta_{\mathbf{I}}}{\|x_{2,0}\|^2}\left(\frac{L_f}{\delta\beta_{\mathbf{I}}}e^{\alpha t}\big(\beta_{\mathbf{I}}Ce^{\delta t}+1\big)\big(e^{\delta t}-1\big) + \frac{1}{\beta_{\mathbf{I}}^2}\big(e^{\alpha t}-1\big)\right)\|x_{2,0}\|^2 \\
&\quad + \frac{\beta_{\mathbf{I}}}{\|x_{2,0}\|^2}\left(\frac{H_f}{\delta}e^{(\alpha+2\delta)t}\|x_{2,0}\|\cdot\big(e^{\delta t}-1\big)\right. \\
&\quad \left. + \mathscr{L}_f\big(e^{(\alpha+\delta)t}+e^{\alpha t}\big)\big(e^{\delta t}-1\big) + \mathscr{L}_f\big(e^{\alpha t}-1\big)\right)\|x_{2,0}\|^2 \\
&\overset{\text{(ii)}}{\leq} \frac{L_f}{\delta}e^{\delta t}(\beta_{\mathbf{I}}Ce^{\delta t}+1)(e^{\delta t}-1) + \frac{1}{\beta_{\mathbf{I}}}(e^{\delta t}-1) \\
&\quad + \beta_{\mathbf{I}}\Big(\beta_{\mathbf{I}}H_f\delta^{-1}\cdot e^{3\delta t}\big\|\nabla f(x_{1,0})\big\|\cdot\big(e^{\delta t}-1\big) \\
&\quad + \mathscr{L}_f\big(e^{2\delta t}+e^{\delta t}\big)\big(e^{\delta t}-1\big) + \mathscr{L}_f\big(e^{\delta t}-1\big)\Big) \\
&= \Big(L_f\delta^{-1}\cdot e^{\delta t}(\beta_{\mathbf{I}}Ce^{\delta t}+1) + \frac{1}{\beta_{\mathbf{I}}} + \frac{\beta_{\mathbf{I}}^2H_f}{\delta}e^{3\delta t}\big\|\nabla f(x_{1,0})\big\| \\
&\quad + \beta_{\mathbf{I}}\mathscr{L}_f(e^{2\delta t}+e^{\delta t}) + \beta_{\mathbf{I}}\mathscr{L}_f\Big)\big(e^{\delta t}-1\big) \\
&=: \phi\Big(t, \big\|\nabla f(x_{1,0})\big\|\Big),
\end{aligned}
$$

where the inequality (i) follows from the implications of Steps 2 and 3, and the equality (ii) is an immediate consequence of the relation $\alpha<\delta$ and the equality $x_{2,0}=-\beta_{\mathbf{I}}\nabla f(x_{1,0})$.

Step 5: Consider a_1 defined in (2.8d) and recall that $u_{\mathbf{I}}(x_{,0})$ by design lies inside the input interval $[\underline{u}_{\mathbf{I}}, \overline{u}_{\mathbf{I}}]$. The quantity a_1 is a lower bound on the distance of $u_{\mathbf{I}}(x_{,0})$ to the

boundaries of the interval $[\underline{u}_{\mathbf{I}}, \overline{u}_{\mathbf{I}}]$. Thus, the inter-jump interval $\tau_{\mathbf{I}}$ satisfies

$$\tau_{\mathbf{I}} \geq \max\{t \geq 0 : \ |u_{\mathbf{I}}(x) - u_{\mathbf{I}}(x_{,0})| \leq a_1\} \geq \max\Big\{t \geq 0 : \ \phi\Big(t, \big\|\nabla f(x_{1,0})\big\|\Big) \leq a_1\Big\},$$

where the second inequality is implied by the analysis provided in Step 4. Consider a positive constant $r > 1$. One can infer for every $t \in [0, \delta^{-1}\log r]$ that

$$\begin{aligned}\phi\Big(t, \big\|\nabla f(x_{1,0})\big\|\Big) &\leq \Big(rL_f\delta^{-1}(r\beta_{\mathbf{I}}C + 1) + \beta_{\mathbf{I}}^{-1} + r^3\beta_{\mathbf{I}}^2 H_f\delta^{-1}\big\|\nabla f(x_{1,0})\big\| \\ &\quad + (r^2 + r)\beta_{\mathbf{I}}\mathscr{L}_f + \beta_{\mathbf{I}}\mathscr{L}_f\Big)(e^{\delta t} - 1) \\ &= \Big(a_2 + a_3\big\|\nabla f(x_{1,0})\big\|\Big)(e^{\delta t} - 1) \\ &=: \phi'\Big(t, \big\|\nabla f(x_{1,0})\big\|\Big),\end{aligned}$$

where the constants a_2 and a_3 are defined in (2.8e), (2.8f), respectively, and the inequality $e^{\delta t} < r$ is used. Suppose now τ' is the lower bound of the inter jump in (2.7). Then $\phi'\Big(\tau', \big\|\nabla f(x_{1,0})\big\|\Big) = a_1$, where the constant a_1 is defined in (2.8d). It is straightforward to establish the assertion made in (2.7).

In the second part of the assertion, we should show that the proposed lower bound in (2.7) is uniformly away from zero along any trajectories of the hybrid system. To this end, we only need to focus on the term $\|\nabla f\big(x_1(t)\big)\|$. Recall that Theorem 2.3.1 effectively implies that $\lim_{t\to\infty} \|\nabla f\big(x_1(t)\big)\| = 0$, possibly not in a monotone manner though. This observation allows us to deduce that $M := \sup_{t\geq 0} \|\nabla f\big(x_1(t)\big)\| < \infty$. Using the uniform bound M, we have a minimum non-zero inter-jump interval, giving rise to a Zeno-free behavior for all solution trajectories.

2.4.3. PROOF OF THEOREM 2.3.4

The proof follows a similar idea as in Theorem 2.3.1 but the required technical steps are somewhat different, leading to another set of technical assumptions. In the first step, we begin with describing on how the chosen input $u_{\mathbf{II}}(x)$ in (2.9b) ensures achieving the desired exponential convergence rate $\mathscr{O}\big(e^{-\alpha t}\big)$. Let us define the set $\mathscr{E}_\alpha := \Big\{x \in \mathbb{R}^{2n} : \ \alpha\big(f(x_1) - f^*\big) < \langle \nabla f(x_1), -x_2\rangle\Big\}$. We demonstrate that as long as a solution trajectory of the continuous flow (2.9a) is contained in the set $\mathscr{E}_\alpha$, the function f obeys the exponential decay (2.5). To this end, observe that if $\big(x_1(t), x_2(t)\big) \in \mathscr{E}_\alpha$,

$$\frac{d}{dt}\Big(f\big(x_1(t)\big) - f^*\Big) = \big\langle \nabla f\big(x_1(t)\big), x_2(t)\big\rangle \leq -\alpha\big(f(x_1) - f^*\big).$$

The direct application of Gronwall's inequality, see [44, Lemma A.1], to the above inequality yields the desired convergence claim (2.5). Hence, it remains to guarantee that the solution trajectory renders the set $\mathscr{E}_\alpha$ invariant. Let us define the quantity

$$\sigma(t) := \langle \nabla f\big(x_1(t)\big), x_2(t)\rangle + \alpha\Big(f\big(x_1(t)\big) - f^*\Big).$$

By construction, if $\sigma(t) < 0$, it follows that $\big(x_1(t), x_2(t)\big) \in \mathscr{E}_\alpha$. As a result, if we synthesize the feedback input $u_{\mathbf{II}}(x)$ such that $\dot{\sigma}(t) \leq 0$ along the solution trajectory of (2.9a), the

value of $\sigma(t)$ does not increase, and as such

$$\big(x_1(t), x_2(t)\big) \in \mathscr{E}_\alpha,\ \forall\, t \geq 0 \iff \big(x_1(0), x_2(0)\big) \in \mathscr{E}_\alpha.$$

To ensure non-positivity property of $\dot{\sigma}(t)$, note that we have

$$\begin{aligned}
\dot{\sigma}(x) &= \langle \nabla^2 f(x_1)x_2, x_2\rangle + \langle \nabla f(x_1), \dot{x}_2\rangle + \alpha\langle \nabla f(x_1), x_2\rangle \\
&= \langle \nabla^2 f(x_1)x_2, x_2\rangle + \langle \nabla f(x_1), -x_2 - u_{\mathbf{II}}(x)\nabla f(x_1)\rangle + \alpha\langle \nabla f(x_1), x_2\rangle \\
&= \langle \nabla^2 f(x_1)x_2, x_2\rangle + \langle \nabla f(x_1), -x_2\rangle - u_{\mathbf{II}}(x)\|\nabla f(x_1)\|^2 - \alpha\langle \nabla f(x_1), -x_2\rangle \\
&= \langle \nabla^2 f(x_1)x_2, x_2\rangle + (1-\alpha)\langle \nabla f(x_1), -x_2\rangle - u_{\mathbf{II}}(x)\|\nabla f(x_1)\|^2 = 0,
\end{aligned}$$

where the last equality follows from the definition of the proposed control law (2.9b). It is worth noting that one can simply replace the information of the Hessian $\nabla^2 f\big(x_1(t)\big)$ with the upper bound L_f and still arrive at the desired inequality, see also Remark 2.3.9 with regards to the 1st-order information oracle. Up to now, we showed that the structure of the control feedback guarantees the α-exponential convergence. It then remains to ensure that $x(0) \in \mathscr{E}_\alpha$. Consider the initial state $x_2(0) = -\beta_{\mathrm{II}}\nabla f\big(x_1(0)\big)$. Notice that

$$\begin{aligned}
\alpha\Big(f\big(x_1(0)\big) - f^*\Big) &\leq \frac{\alpha}{2\mu_f}\big\|\nabla f\big(x_1(0)\big)\big\|^2 \\
&= \frac{\alpha}{2\mu_f\beta_{\mathrm{II}}}\langle -x_2(0), \nabla f\big(x_1(0)\big)\rangle \\
&\leq \langle \nabla f\big(x_1(0)\big), -x_2(0)\rangle,
\end{aligned}$$

where in the first line we use (A2), and in the last line the condition (2.10c) is employed. Suppose $\big(x_1^\top(0), x_2^\top(0)\big)^\top$ as the jump state x^+. It is evident that the range space of the jump map (2.9d) lies inside the set $\mathscr{E}_\alpha$. At last, it is required to show that the jump policy is well-defined in the sense that the trajectory lands in the interior of the flow set $\mathscr{C}_{\mathbf{I}}$ (2.9c), i.e., the control values also belong to the admissible set $[\underline{u}_{\mathbf{II}}, \overline{u}_{\mathbf{II}}]$. To this end, we only need to take into account the initial control value since the switching law is continuous in the states and serves the purpose by design. Suppose that $x^+ \in \mathscr{C}_{\mathbf{II}}$, we then have the sufficient requirements

$$\begin{aligned}
\underline{u}_{\mathbf{II}} < \frac{-\ell_f\beta_{\mathrm{II}}^2\|\nabla f(x_1^+)\|^2 + (1-\alpha)\beta_{\mathrm{II}}\|\nabla f(x_1^+)\|^2}{\|\nabla f(x_1^+)\|^2} \\
\leq u_{\mathbf{II}}(x^+) \leq \\
\frac{L_f\beta_{\mathrm{II}}^2\|\nabla f(x_1^+)\|^2 + (1-\alpha)\beta_{\mathrm{II}}\|\nabla f(x_1^+)\|^2}{\|\nabla f(x_1^+)\|^2} < \overline{u}_{\mathbf{II}},
\end{aligned}$$

where the relations (2.9b) and (A1) are considered. Factoring out the term $\|\nabla f(x_1^+)\|^2$ leads to the sufficiency requirements given in (2.10a) and (2.10b). Hence, the claim of Theorem 2.3.4 follows.

2.4.4. PROOF OF THEOREM 2.3.5

In order to facilitate the argument regarding the proof of Theorem 2.3.5, we begin with providing a lemma describing the 2-norm behaviors of $\langle \nabla f(x_1), -x_2\rangle$, x_2, and $\nabla f(x_1)$. For the sake of brevity, we employ the same notations used in Subsection 2.4.2, as well.

Lemma 2.4.3 (Growth bounds)**.** *Consider the continuous-time hybrid control system $(\mathcal{H})$ with the respective parameters (2.9) satisfying (2.10) where the function f satisfies Assumptions (A1) and (A2). Suppose the control system is initiated from $\big(x_{1,0}, \beta_{\boldsymbol{II}}\nabla f(x_{1,0})\big)$ for some $x_{1,0} \in \mathbb{R}^n$. Then,*

$$\langle \nabla f(x_1), -x_2\rangle = \beta_{\boldsymbol{II}} e^{-\alpha t}\|\nabla f(x_{1,0})\|^2, \tag{2.23a}$$

$$\|x_2\| \leq D(t)\|\nabla f(x_{1,0})\|, \tag{2.23b}$$

$$\underline{\eta}(t)\|\nabla f(x_{1,0})\| \leq \|\nabla f(x_1)\| \leq \overline{\eta}(t)\|\nabla f(x_{1,0})\|, \tag{2.23c}$$

with the time-varying scalars D, $\underline{\eta}$, and $\overline{\eta}$ given by

$$D(t) := \Big(\beta_{\boldsymbol{II}}^2 e^{-2t} + \beta_{\boldsymbol{II}} U\big(1 - e^{-2t}\big)\Big)^{\frac{1}{2}}, \tag{2.24a}$$

$$\underline{\eta}(t) := 1 - \mathscr{L}_f(\beta_{\boldsymbol{II}}^2 + \beta_{\boldsymbol{II}} U)^{\frac{1}{2}}\, t, \tag{2.24b}$$

$$\overline{\eta}(t) := 1 + \mathscr{L}_f(\beta_{\boldsymbol{II}}^2 + \beta_{\boldsymbol{II}} U)^{\frac{1}{2}}\, t, \tag{2.24c}$$

respectively, where $U := \max\{\overline{u}_{\boldsymbol{II}}, -\underline{u}_{\boldsymbol{II}}\}$ and $\mathscr{L}_f := \max\{\ell_f, L_f\}$.

Proof. Considering the flow dynamics (2.9a) and the feedback input (2.9b), one obtains

$$\begin{aligned}
\frac{d}{dt}\langle \nabla f(x_1), -x_2\rangle &= \langle \nabla^2 f(x_1)x_2, -x_2\rangle + \langle \nabla f(x_1), -\dot{x}_2\rangle \\
&= \langle \nabla^2 f(x_1)x_2, -x_2\rangle + \langle \nabla f(x_1), x_2 + u_{\mathbf{II}}(x)\nabla f(x_1)\rangle \\
&= \langle \nabla^2 f(x_1)x_2, -x_2\rangle + \langle \nabla f(x_1), x_2\rangle + u_{\mathbf{II}}(x)\|\nabla f(x_1)\|^2 \\
&= \langle \nabla^2 f(x_1)x_2, -x_2\rangle + \langle \nabla f(x_1), x_2\rangle + \langle \nabla^2 f(x_1)x_2, x_2\rangle - (1-\alpha)\langle \nabla f(x_1), x_2\rangle \\
&= -\alpha\langle \nabla f(x_1), -x_2\rangle,
\end{aligned}$$

and as a result given the initial state $\big(x_{1,0}, -\beta_{\mathbf{II}}\nabla f(x_{1,0})\big)$, the equality given in (2.23a) is valid. We next turn to establish that (2.23b) holds. Let us define $h(t) = \|x_2\|^2$. Hence,

$$\begin{aligned}
\frac{d}{dt}h(t) &\overset{\text{(i)}}{=} 2\langle x_2, -x_2 - u_{\mathbf{II}}(x)\nabla f(x_1)\rangle \\
&= -2\|x_2\|^2 + 2u_{\mathbf{II}}(x)\langle \nabla f(x_1), -x_2\rangle \\
&\overset{\text{(ii)}}{=} -2h(t) + 2u_{\mathbf{II}}(x)\beta_{\mathbf{II}} e^{-\alpha t}\|\nabla f(x_{1,0})\|^2 \\
&\leq -2h(t) + 2U\beta_{\mathbf{II}}\|\nabla f(x_{1,0})\|^2,
\end{aligned}$$

where we made use of the flow dynamics (2.9a) in the inequality (i) and the equation (2.23a) in the equality (ii). We then use the Gronwall's inequality to infer that

$$\begin{aligned}
\|x_2\|^2 &\leq e^{-2t}\|x_{2,0}\|^2 + \int_0^t e^{-2(t-\tau)} 2U\beta_{\mathbf{II}}\big\|\nabla f(x_{1,0})\big\|^2 d\tau \\
&= e^{-2t}\beta_{\mathbf{II}}^2\big\|\nabla f(x_{1,0})\big\|^2 + e^{-2t}2U\beta_{\mathbf{II}}\big\|\nabla f(x_{1,0})\big\|^2 \int_0^t e^{2\tau} d\tau \\
&= e^{-2t}\big\|\nabla f(x_{1,0})\big\|^2\Big(\beta_{\mathbf{II}}^2 e^{-2t} + \beta_{\mathbf{II}} U\big(1 - e^{-2t}\big)\Big) \\
&=: D^2(t)\big\|\nabla f(x_{1,0})\big\|^2,
\end{aligned}$$

where $D(t)$ is defined in (2.24a). As a result, the claim in (2.23b) holds. The argument to show the last claim in Lemma 2.4.3 is discussed now. Let us define $g(t) := \|\nabla f(x_1)\|^2$. Observe that

$$\frac{d}{dt}g(t) = 2\langle \nabla^2 f(x_1)x_2, \nabla f(x_1)\rangle,$$

and as a result

$$\left|\frac{d}{dt}g(t)\right| \overset{\text{(i)}}{\leq} 2\mathscr{L}_f\|x_2\| \cdot \|\nabla f(x_1)\| = 2\mathscr{L}_f\|x_2\|\sqrt{g(t)} \overset{\text{(ii)}}{\leq} 2\mathscr{L}_f D(t)\|\nabla f(x_{1,0})\|\sqrt{g(t)},$$

where the inequalities (i) and (ii) are implied by Assumption (A1) and the inequality (2.23b), respectively. Hence, we deduce that

$$\frac{d}{dt}g(t) \geq -2\mathscr{L}_f D(t)\|\nabla f(x_{1,0})\|\sqrt{g(t)},$$

and as a consequence

$$\frac{dg(t)}{\sqrt{g(t)}} \geq -2\mathscr{L}_f D(t)\|\nabla f(x_{1,0})\|dt.$$

Integrating the two sides of the above inequality results in

$$\begin{aligned}
\sqrt{g(t)} - \sqrt{g(0)} &\geq -\mathscr{L}_f\|\nabla f(x_{1,0})\|\int_0^t D(\tau)d\tau \\
&= -\mathscr{L}_f\|\nabla f(x_{1,0})\|\int_0^t \left(\beta_{\mathbf{II}}^2 e^{-2\tau} + \beta_{\mathbf{II}}U\big(1-e^{-2\tau}\big)\right)^{\frac{1}{2}} d\tau \\
&\geq -\mathscr{L}_f\|\nabla f(x_{1,0})\|\int_0^t \left(\beta_{\mathbf{II}}^2 + \beta_{\mathbf{II}}U\right)^{\frac{1}{2}} d\tau \\
&= -\mathscr{L}_f\|\nabla f(x_{1,0})\|\left(\beta_{\mathbf{II}}^2 + \beta_{\mathbf{II}}U\right)^{\frac{1}{2}} t.
\end{aligned}$$

Based on the above analysis and the definition of $g(t)$, it follows that

$$\|\nabla f(x_1)\| \geq \underline{\eta}(t)\|\nabla f(x_{1,0})\|,$$

where $\underline{\eta}(t)$ is given in (2.24b). Proceeding with a similar approach to the one presented above, one can use the inequality

$$\frac{d}{dt}g(t) \leq 2\mathscr{L}_f D(t)\|\nabla f(x_{1,0})\|\sqrt{g(t)},$$

and infer that

$$\|\nabla f(x_1)\| \leq \overline{\eta}(t)\|\nabla f(x_{1,0})\|,$$

where $\overline{\eta}(t)$ is defined in (2.24c). Thus, the last claim in Lemma 2.4.3 also holds. □

Proof of Theorem 2.3.5: We are now in a position to formally state the proof of Theorem 2.3.5. Consider the parameter δ as defined in Theorem 2.3.5. Intuitively, this quantity represents a lower bound on the distance of $u_{\mathbf{II}}(0)$ from the endpoints of the flow set interval. Thus, one can obtain a lower bound on the inter-jump interval $\tau_{\mathbf{II}}$ as follows

$$\tau_{\mathbf{II}} \geq \sup\{t>0 : |u_{\mathbf{II}}(t) - u_{\mathbf{II}}(0)| \leq \delta\}. \tag{2.25}$$

On the other hand, given the structure of $u_{\mathbf{II}}$ in (2.9b),

$$-\frac{\ell_f \|x_2\|^2}{\|\nabla f(x_1)\|^2} + (1-\alpha)\frac{\beta_{\mathbf{II}} e^{-\alpha t}\|\nabla f(x_{1,0})\|^2}{\|\nabla f(x_1)\|^2} \leq u_{\mathbf{II}}(t) \leq \frac{L_f \|x_2\|^2}{\|\nabla f(x_1)\|^2} + (1-\alpha)\frac{\beta_{\mathbf{II}} e^{-\alpha t}\|\nabla f(x_{1,0})\|^2}{\|\nabla f(x_1)\|^2},$$

since the function f satisfies Assumption (A1). In light of Lemma 2.4.3 and considering the above relation, one can infer that for $\alpha \leq 1$, we name Case(i),

$$\underline{e}(t) := -\frac{\ell_f D(t)^2}{\underline{\eta}(t)^2} + (1-\alpha)\frac{\beta_{\mathbf{II}} e^{-\alpha t}}{\overline{\eta}(t)^2} \leq u_{\mathbf{II}}(t) \leq \frac{L_f D(t)^2}{\underline{\eta}(t)^2} + (1-\alpha)\frac{\beta_{\mathbf{II}} e^{-\alpha t}}{\underline{\eta}(t)^2} =: \overline{e}(t), \tag{2.26a}$$

and that for $\alpha > 1$, we denote by Case (ii),

$$\underline{p}(t) := -\frac{\ell_f D(t)^2}{\underline{\eta}(t)^2} + (1-\alpha)\frac{\beta_{\mathbf{II}} e^{-\alpha t}}{\underline{\eta}(t)^2} \leq u_{\mathbf{II}}(t) \leq \frac{L_f D(t)^2}{\underline{\eta}(t)^2} + (1-\alpha)\frac{\beta_{\mathbf{II}} e^{-\alpha t}}{\overline{\eta}(t)^2} =: \overline{p}(t). \tag{2.26b}$$

According to the above discussion, we employ (2.26) to obtain a lower bound $\tau_{\mathbf{II}}$ instead of using (2.25). Consider a time instant $t_\circ$ such that $t_\circ < 1/\omega$ where ω is defined in Theorem 2.3.5.

Case (i) ($\alpha \leq 1$): Let us denote $\sup_{t\in[0,t_\circ]} \dot{\overline{e}}(t)$ by b_1. Observe that

$$\begin{aligned}
\dot{\overline{e}}(t) &= \frac{2L_f\beta_{\mathbf{II}}e^{-2t}(-\beta_{\mathbf{II}}+U)(1-\omega t)^2 + 2\omega(1-\omega t)L_f\beta_{\mathbf{II}}\big(\beta_{\mathbf{II}}e^{-2t}+U(1-e^{-2t})\big)}{(1-\omega t)^4} \\
&+ (1-\alpha)\frac{-\alpha\beta_{\mathbf{II}}e^{-\alpha t}(1-\omega t)^2 + 2\omega(1-\omega t)\beta_{\mathbf{II}}e^{-2t}}{(1-\omega t)^4} \\
&\leq \frac{2L_f\beta_{\mathbf{II}}Ue^{-2t}(1-\omega t)^2 + 2\omega(1-\omega t)L_f\beta_{\mathbf{II}}\big(\beta_{\mathbf{II}}e^{-2t}+U(1-e^{-2t})\big)}{(1-\omega t)^4} \\
&+ (1-\alpha)\frac{2\omega(1-\omega t)\beta_{\mathbf{II}}e^{-2t}}{(1-\omega t)^4} \\
&\leq \frac{2L_f\beta_{\mathbf{II}}\big(U+\omega(\beta_{\mathbf{II}}+U)\big)}{(1-\omega t)^3} + (1-\alpha)\frac{2\omega\beta_{\mathbf{II}}}{(1-\omega t)^3} \\
&\leq \frac{2L_f\beta_{\mathbf{II}}\big(U+\omega(\beta_{\mathbf{II}}+U)\big)}{(1-\omega t_\circ)^3} + (1-\alpha)\frac{2\omega\beta_{\mathbf{II}}}{(1-\omega t_\circ)^3} =: b_1,
\end{aligned}$$

considering (2.26a). Hence, $\overline{e}(t) \leq b_1 t + \overline{e}(0)$ and as a result

$$\tau_{\mathbf{II}} \geq \max\{t \in (0,t_\circ] :\ b_1 t + \overline{e}(0) - \overline{e}(0) \leq \delta\} = \min\{t_\circ, \delta/b_1\}, \tag{2.27}$$

by virtue of the fact that $b_1 t + \overline{e}(0)$ is a monotonically increasing function that upper bounds $u_{\mathbf{II}}(t)$. Now, let us define $b_2 := \inf_{t\in(0,t_\circ]} \dot{\underline{e}}(t)$. Notice that

$$\begin{aligned}
\dot{\underline{e}}(t) &= \frac{2\ell_f\beta_{\mathbf{II}}e^{-2t}(\beta_{\mathbf{II}}-U)(1-\omega t)^2 - 2\omega(1-\omega t)\ell_f\beta_{\mathbf{II}}\big(\beta_{\mathbf{II}}e^{-2t}+U(1-e^{-2t})\big)}{(1-\omega t)^4} \\
&+ (1-\alpha)\frac{-\alpha\beta_{\mathbf{II}}e^{-\alpha t}(1+\omega t)^2 - 2\omega(1+\omega t)\beta_{\mathbf{II}}e^{-2t}}{(1+\omega t)^4} \\
&\geq \frac{-2\ell_f\beta_{\mathbf{II}}e^{-2t}U(1-\omega t)^2 - 2\omega(1-\omega t)\ell_f\beta_{\mathbf{II}}\big(\beta_{\mathbf{II}}e^{-2t}+U(1-e^{-2t})\big)}{(1-\omega t)^4} \\
&- (1-\alpha)\frac{\alpha\beta_{\mathbf{II}}e^{-\alpha t}(1+\omega t)^2 + 2\omega(1+\omega t)\beta_{\mathbf{II}}e^{-2t}}{(1+\omega t)^4} \\
&\geq -\frac{2\ell_f\beta_{\mathbf{II}}\big(U+\omega(\beta_{\mathbf{II}}+U)\big)}{(1-\omega t_\circ)^3} - (1-\alpha)\frac{\alpha\beta_{\mathbf{II}}(1+\omega t_\circ)+2\omega\beta_{\mathbf{II}}}{1} =: -b_2.
\end{aligned}$$

Thus, $\underline{e}(t) \geq -b_2 t + \underline{e}(0)$ and as a consequence

$$\tau_{\mathbf{II}} \geq \max\{t \in (0,t_\circ] :\ \underline{e}(0) - \big(-b_2 t + \underline{e}(0)\big) \leq \delta\} = \min\{t_\circ, \delta/b_2\}, \tag{2.28}$$

because the function $-b_2 t + \underline{e}(0)$ is a monotonically decreasing function that lower bounds $u_{\mathbf{II}}(t)$.

Case (ii) ($\alpha > 1$): Much of this case follows the same line of reasoning used in Case (i). We thus provide only main mathematical derivations and refer the reader to the previous case for the argumentation. Define $b_3 := \sup_{t\in(0,t_\circ]} \dot{\overline{p}}(t)$. One can deduce from (2.26b)

that

$$\begin{aligned}\dot{\overline{p}}(t) &= \frac{2L_f\beta_{\mathbf{II}}e^{-2t}(-\beta_{\mathbf{II}}+U)(1-\omega t)^2+2\omega(1-\omega t)L_f\beta_{\mathbf{II}}\big(\beta_{\mathbf{II}}e^{-2t}+U(1-e^{-2t})\big)}{(1-\omega t)^4}\\ &+(1-\alpha)\frac{-\alpha\beta_{\mathbf{II}}e^{-\alpha t}(1+\omega t)^2-2\omega(1+\omega t)\beta_{\mathbf{II}}e^{-2t}}{(1+\omega t)^4}\\ &\le \frac{2L_f\beta_{\mathbf{II}}\big(U+\omega(\beta_{\mathbf{II}}+U)\big)}{(1-\omega t_\circ)^3}+(\alpha-1)\frac{\alpha\beta_{\mathbf{II}}(1+\omega t_\circ)+2\omega\beta_{\mathbf{II}}}{1}=:b_3.\end{aligned}$$

Hence, $\overline{p}(t)\le b_4 t+\overline{p}(0)$ and as a result

$$\tau \ge \min\{t_\circ, \delta/b_3\}. \tag{2.29}$$

Finally, define $\dot{\underline{p}}(t) := \inf_{t\in(0,t_\circ]}\underline{p}(t)$ from which it follows that

$$\begin{aligned}\dot{\underline{p}}(t) &= \frac{2\ell_f\beta_{\mathbf{II}}e^{-2t}(\beta_{\mathbf{II}}-U)(1-\omega t)^2-2\omega(1-\omega t)\ell_f\beta_{\mathbf{II}}\big(\beta_{\mathbf{II}}e^{-2t}+U(1-e^{-2t})\big)}{(1-\omega t)^4}\\ &+(1-\alpha)\frac{-\alpha\beta_{\mathbf{II}}e^{-\alpha t}(1-\omega t)^2+2\omega(1-\omega t)\beta_{\mathbf{II}}e^{-2t}}{(1-\omega t)^4}\\ &\ge -\frac{2\ell_f\beta_{\mathbf{II}}\big(U+\omega(\beta_{\mathbf{II}}+U)\big)}{(1-\omega t_\circ)^3}-(\alpha-1)\frac{2\omega\beta_{\mathbf{II}}}{(1-\omega t_\circ)^3}=:-b_4,\end{aligned}$$

considering (2.26b). Now, since $\underline{p}(t)\ge -b_4 t+\underline{p}(0)$, it is implied that

$$\tau_{\mathbf{II}} \ge \min\{t_\circ, \delta/b_4\}. \tag{2.30}$$

Notice that based on the relations derived in (2.28)-(2.30),

$$\tau_{\mathbf{II}} \ge \min\Big\{t_\circ, \frac{2\mathscr{L}_f\beta_{\mathbf{II}}\big(U+\omega(\beta_{\mathbf{II}}+U)\big)}{(1-\omega t_\circ)^3}+|\alpha-1|\frac{2\omega\beta_{\mathbf{II}}}{(1-\omega t_\circ)^3}+|\alpha-1|\alpha\beta_{\mathbf{II}}(1+\omega t_\circ)\Big\}.$$

Suppose now for some scalar $r\in(0,1)$, $t_\circ$ is chosen such that $t_\circ\le\frac{r}{\omega}$. It is evident that

$$\tau_{\mathbf{II}} \ge \min\Big\{\frac{r}{\omega}, \delta\Big/\Big(\frac{2\mathscr{L}_f\beta_{\mathbf{II}}\big(U+\omega(\beta_{\mathbf{II}}+U)\big)}{(1-r)^3}+|\alpha-1|\frac{2\omega\beta_{\mathbf{II}}}{(1-r)^3}+|\alpha-1|\alpha\beta_{\mathbf{II}}(1+r)\Big)\Big\}.$$

It turns out that the relation (2.11) in Theorem 2.3.5 is valid and this concludes the proof.

2.4.5. PROOF OF THEOREM 2.3.11

In what follows, we provide the proof for the structure **II** and refer the interested reader to [1, Theorem 3.7] for the structure **I**. We emphasize that the technical steps to establish a stable discretization for both structures are similar.

According to the forward-Euler method, the velocity $\dot{x}_1$ and the acceleration $\dot{x}_2$ in the dynamics ($\mathscr{H}$) with (2.9) are discretized as follows:

$$\frac{x_1^{k+1}-x_1^k}{s} = x_2^k, \tag{2.31a}$$

$$\frac{x_2^{k+1}-x_2^k}{s} = -u_{d,\mathbf{II}}(x^k)\nabla f(x_1^k)-x_2^k, \tag{2.31b}$$

where the discrete input $u_{d,\mathbf{II}}(x^k) = u_{\mathbf{II}}(x^k)$. Now, observe that the definition of the flow set $\mathscr{C}_{d,\mathbf{II}}$ (2.13b) implies

$$c_1\|x_2^k\|^2 \leq \|\nabla f(x_1^k)\|^2 \leq c_2\langle \nabla f(x_1^k), -x_2^k\rangle \leq c_2\|\nabla f(x_1^k)\|\cdot\|x_2^k\|,$$

where the extra inequality follows from the Cauchy-Schwarz inequality ($\forall\ a,b\in\mathbb{R}^n$, $\langle a,b\rangle \leq \|a\|\cdot\|b\|$). In order to guarantee that the flow set $\mathscr{C}_{d,\mathbf{II}}$ is non-empty the relation (2.16a) should hold between the parameters c_1 and c_2 since $\sqrt{c_1} \leq \frac{\|\nabla f(x_1^k)\|}{\|x_2^k\|} \leq c_2$. Suppose next that the parameters c_1, c_2, and β satisfy (2.16b). Multiplying (2.16b) by $\|\nabla f(x_1^k)\|$, observe that the range space of the jump map $G_{d,\mathbf{II}}(x^k) = \big((x^k)^\top, -\beta\nabla^\top f(x^k)\big)^\top$ is inside the flow set $\mathscr{C}_{d,\mathbf{II}}$ (2.13b). From the fact that the discrete dynamics (2.12) evolves respecting the flow set $\mathscr{C}_{d,\mathbf{II}}$ defined in (2.13b), we deduce

$$\begin{aligned}
f(x_1^{k+1}) - f(x_1^k) \quad &\leq \langle \nabla f(x_1^k), x_1^{k+1} - x_1^k\rangle + \frac{L_f}{2}\|x_1^{k+1} - x_1^k\|^2\\
&\leq -s\langle \nabla f(x_1^k), -x_2^k\rangle + \frac{L_f s^2}{2}\|x_2^k\|^2\\
&< -\frac{s}{c_2}\|\nabla f(x_1^k)\|^2 + \frac{L_f s^2}{2c_1}\|\nabla f(x_1^k)\|^2\\
&= \big(-\frac{s}{c_2} + \frac{L_f}{2c_1}s^2\big)\|\nabla f(x_1^k)\|^2\\
&\leq 2\mu_f\big(-\frac{s}{c_2} + \frac{L_f}{2c_1}s^2\big)\big(f(x_1^k) - f^*\big),
\end{aligned}$$

where we made use of the relation (2.3), the definition (2.31a), the relation (2.13b), and the assumption (A2), respectively. Then, considering the inequality implied by the first and last terms given above and adding $f(x_1^k) - f^*$ to both sides of the considered inequality, we arrive at

$$f(x_1^{k+1}) - f^* \leq \lambda(s,c_1,c_2,\beta)\Big(f(x_1^k) - f^*\Big)$$

where $\lambda(s,c_1,c_2,\beta)$ is given by (2.15). As a result, if the step size s is chosen such that $s < \frac{2c_1}{c_2 L_f}$ then $\lambda(s,c_1,c_2,\beta) \in (0,1)$. The claim of Theorem 2.3.11 follows.

2.5. NUMERICAL EXAMPLES

In this section a numerical example illustrating the results in this chapter is represented. The example is a least mean square error (LMSE) problem $f(x_1) = \|Ax_1 - b\|^2$ where $x \in \mathbb{R}^5$ denotes the decision variable, $A \in \mathbb{R}^{50\times 5}$ with $L_f = 2\lambda_{\max}(A^\top A) = 136.9832$ and $\mu_f = 2\lambda_{\min}(A^\top A) = 3.6878$, and $b \in \mathbb{R}^{50}$. Since the LMSE function is convex (in our case, this function is strongly convex), we take $\ell_f = 0$. In what follows, we compare the behaviors of the proposed structures **I** and **II** (denoted by **Struct I** and **Struct II**, respectively) along with Nesterov's fast method with the speed restarting scheme proposed by [24] (denoted by **NSR**). We begin with providing the results concerning the continuous-time case. Then, the discrete-time case's results are shown where we employ Algorithm 1 for **Struct I** and **Struct II**.

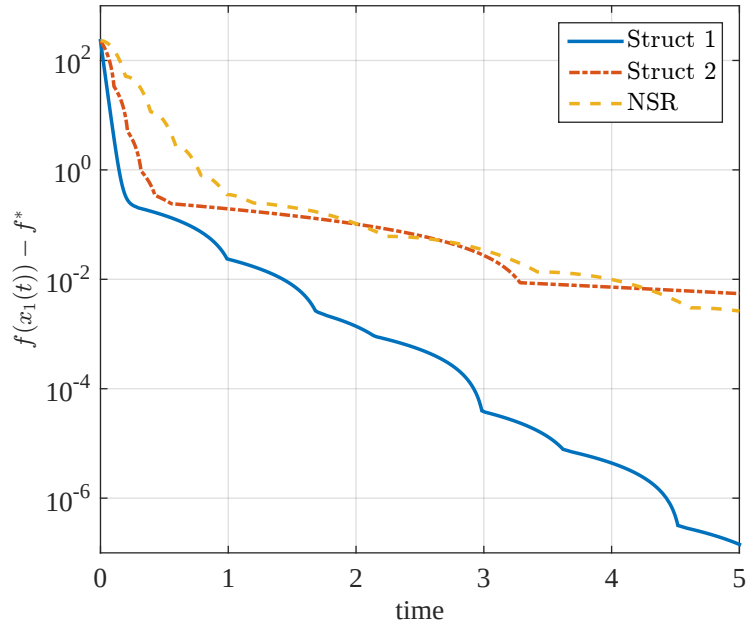
Continuous-time case: The corresponding parameters of **Struct I** and **Struct II** are as follows: $\alpha_I = 0.2$, $\beta_I = 0.1356$, $\underline{u}_{\mathbf{I}} = -14.352$, $\overline{u}_{\mathbf{I}} = 15.1511$; $\alpha_{II} = 0.2$, $\beta_{II} = 0.0298$, $\underline{u}_{\mathbf{II}} = -0.1861$, $\overline{u}_{\mathbf{II}} = 5.7457$. In Figure 2.1a, the behaviors of the suboptimality measure $f\big(x_1(t)\big) - f^*$ of **Struct I**, **Struct II**, and **NSR** are depicted. With regards to Theorem 2.3.2, observe that the length of inter-jump intervals is small during the early stages of simulation. As time progresses and the value of $\nabla f(x_1)$ decreases, the length of inter-jump intervals relatively increases (echoing the same message conveyed in Theorem 2.3.2). The corresponding control inputs are represented in Figure 2.1b. Furthermore, in the case of **Struct I** where $u_{\mathbf{I}}$ plays the role of damping, the input $u_{\mathbf{I}}$ admits a negative range unlike most of the approaches in the literature.

Discrete-time case: Figure 2.2a shows the discrete-time counterparts of the previously mentioned continuous-time dynamics in Figure 2.1. It is evident that the discrete counterparts of our proposed structures perform poorly compared to the NSR's discrete counterpart, reinforcing the assertion of Remark 2.3.12 calling for a smarter discretization technique. The results depicted in Figure 2.2a correspond to the standard parameters involved in each of the algorithm, i.e., the step size $s = 1/L_f$ for the proposed methods in Corollary 2.3.13, and the parameter $k_{\min} = 1$ in NSR. However, these parameters can also be tuned depending on the application at hand. In case of NSR, the role of the parameter $k_{\min}$ is to prevent unnecessary restarting instants that may degrade the overall performance. On the other hand, setting $k_{\min} > 1$ may potentially cause the algorithm to lose its monotonicity property. In Figure 2.2b, we illustrate the best behavior of the three methods with respect to these parameters for this numerical example.

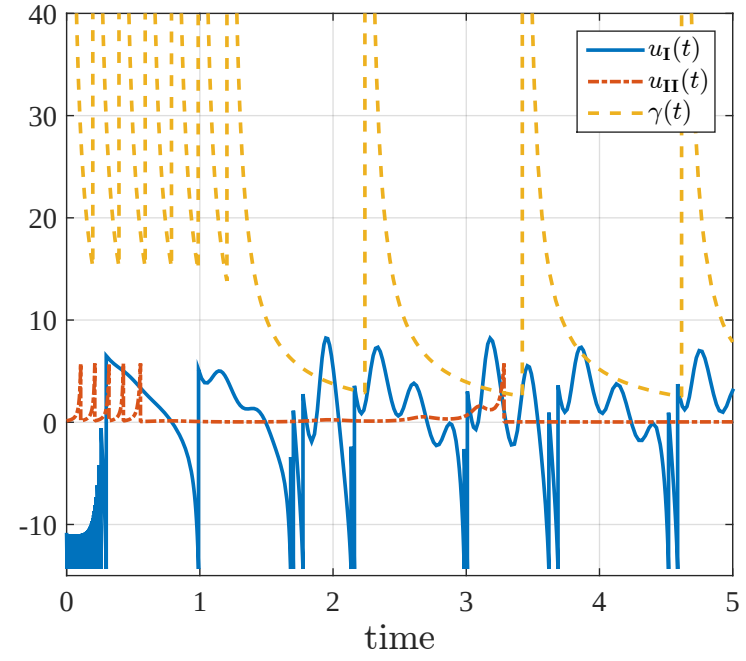
Finally, Figure 2.3a shows how changing $k_{\min}$ affects the performance. The best performance is achieved by setting $k_{\min} = 19$ and the algorithm becomes non-monotonic for $k_{\min} > 19$. With regards to our proposed methods we observe that if one increases the step size s, the performance improves, see Figure 2.3b for **Struct I** and Figure 2.3c for **Struct II**. Moreover, it is obvious that the discrete-time couterparts of **Struct I** and **Struct II** behave in a very similar fashion that has to do with the lack of a proper discretization that can fully exploit the properties of the corresponding feedback input, see Remark 2.3.12.

2.6. CONCLUSIONS

Inspired by a control-oriented viewpoint, we proposed two hybrid dynamical structures to achieve exponential convergence rates for a certain class of unconstrained optimization problems, in a continuous-time setting. The distinctive feature of our methodology is the synthesis of certain inputs in a state-dependent fashion compared to a time-dependent approach followed by most results in the literature. Due to the state dependency of our proposed methods, the time-discretization of continuous-time hybrid dynamical systems is in fact difficult (and to some extent even more involved than the time-varying dynamics that is commonly used in the literature). In this regard, we have been able to show that one can apply the the forward-Euler method to discretize the continuous-time dynamics and still guarantee exponential rate of convergence. Thus, a more in-depth analysis is due. We expect that because of the state dependency of our methods a proper venue to search is geometrical types of discretization.

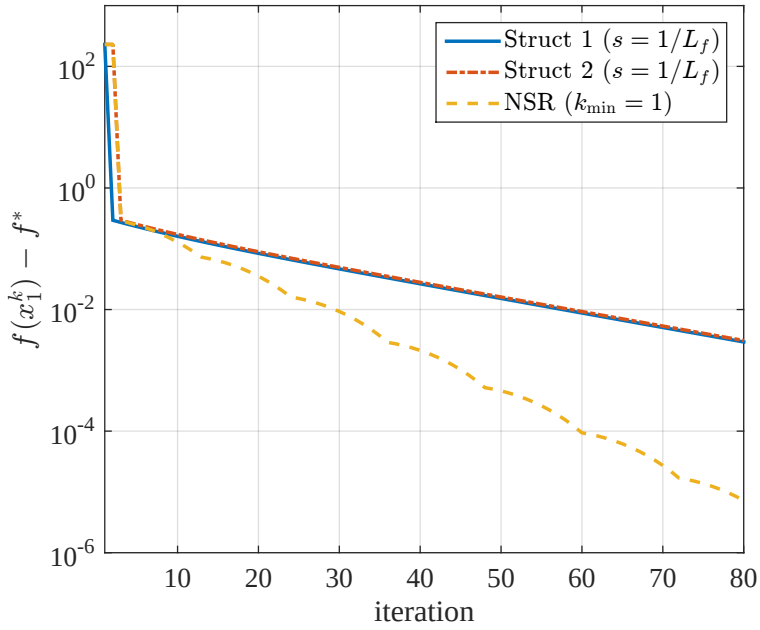


(a) Objective value along system trajectories.

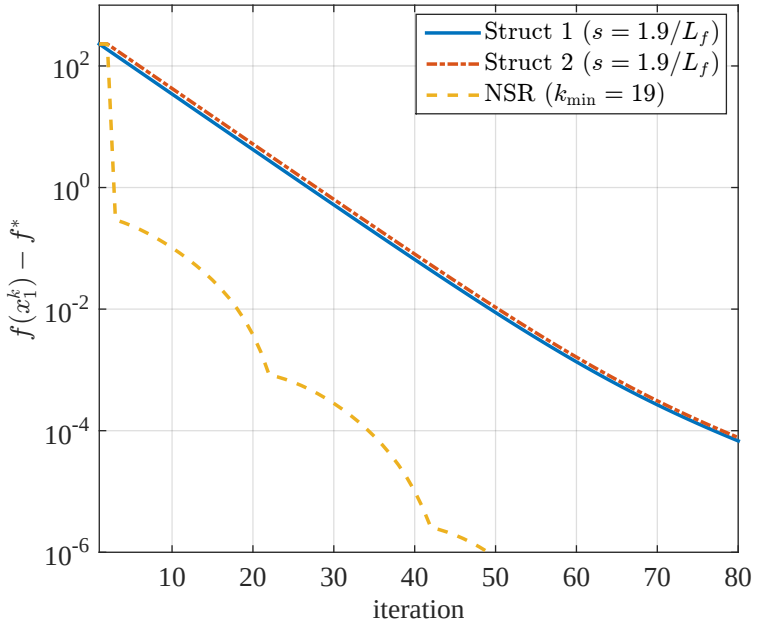


(b) State-dependent and time-varying coefficients.

Figure 2.1: Continuous-time dynamics of **Struct I**, **Struct II**, **NSR**.

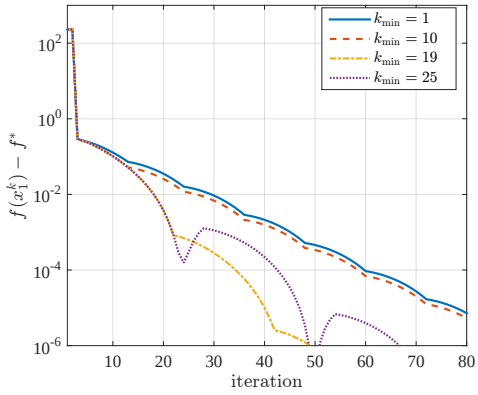


(a) Standard tuning parameters.

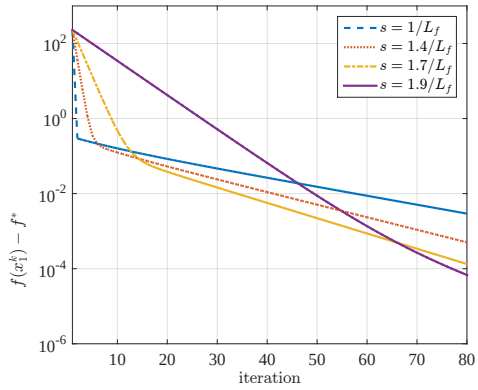


(b) Optimal tuning parameters.

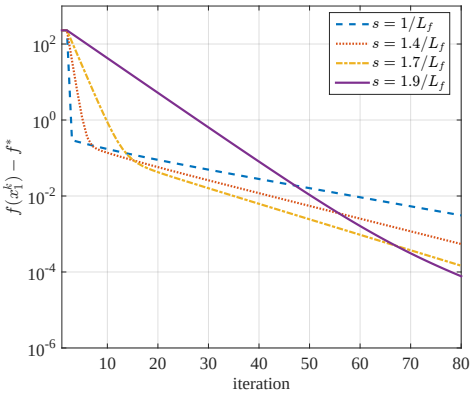
Figure 2.2: Discrete-time dynamics of **Struct I**, **Struct II**, **NSR**.



(a) **NSR**



(b) **Struct I**



(c) **Struct II**

Figure 2.3: Discrete-time dynamics under different tuning parameters.

3

Decentralized Event-Based Policy to Implement an RMPC Approach

In this chapter, we introduce an event-based approach to implement a robust MPC method. The proposed approach is applicable to perturbed, linear time-invariant plants with polytopic constraints. The triggering mechanism is a sequence of hyper-rectangles surrounding the optimal state trajectory. The design of the triggering mechanism is cast as a constrained optimization problem with the volume of each hyper-rectangle as the objective function. We show that the design problem accepts convex nonlinear program and linear program reformulations. On a more practical note, the triggering mechanism suits plants without a centralized sensory node since the triggering mechanism can be evaluated locally at each individual sensor. The materials presented in this chapter are previously reported in [4] and [5].

3.1. INTRODUCTION

Nowadays, applications of networked control systems (NCSs) generally demand an array of compatibility and efficiency measures from control design methods, such as utilization under shared resources, applicability to mobile tasks, and compatibility with digital communication infrastructures. The survey paper [45] and the references therein provides an overview of these emerging challenges. Event-based policies to execute control laws are a class of strategies that aim to systematically address the efficiency problem in the context of communication and computation. In event-based control, the underlying dynamics determine the time at which it is required to update the control action (contrary to the traditional case in which the control action is updated in a *periodic* fashion) [46]. There are two options to implement such an event-based logic in the design: embedded in the sensory system, the so-called *event-triggering control* [47] and [48], or embedded in the controller, the so-called *self-triggering control* [49] and [50]. The responsible entity to determine the update times is known as the *triggering mechanism*. In this regard and at least in a pure theoretical sense, an area of control theory that has witnessed an increased level of interest to employ event-based policies is the class of model predictive control (MPC) methods [51]. There are multiple reasons that encourage such a level of interest ranging from inherent properties of MPC methods to practical advantages. The process of computing the control action is (often) a heavily involved, computational task in MPC methods. One thus hope by employing an event-based implementation to reduce this computational burden. In addition, the predictions that are usually discarded in a standard MPC method can provide one with an object to base the triggering mechanism on.

MPC methods are a class of *on-line* optimization-based control approaches. In a discrete-time setting, the decision variables are the states x_k and the inputs u_k of the considered plant $x_{k+1} = f(x_k, u_k, w_k)$, where w_k denotes possible unknown uncertainties such as exogenous disturbances. (The specific term *robust* MPC (RMPC) is used in the literature when $w_k \neq 0$.) The objective function is a measure of performance we intend to optimize over a finite time horizon N in the future. We base the predictions on an available, *nominal* model $\overline{x}_{k+1} = f(\overline{x}_k, u_k, 0)$ over the horizon, where $\overline{x}_k$ denotes the nominal state. It is evident that a set of constraints originates from the fact that the nominal states and the inputs are subject to the model dynamics. Furthermore, other types of constraints are possibly added due to physical limitations of the underlying plant (such as bounded inputs) or some additional requirements that should be satisfied (such as

the closed-loop stability of the plant or the feasibility of the corresponding optimization problem). The outcomes of the corresponding optimization problem are the input and state trajectories over the horizon (which are referred to as the *optimal* input and state trajectories in the literature). In a standard setting of an MPC method, the corresponding optimization problem is solved at each sampling instant. Then, the first element of the computed, optimal input trajectory is applied to the plant and the remaining entries of the optimal input trajectory are discarded. We refer the interested reader to the survey papers [52] and [53] that discuss about different aspects of MPC. The computational aspect of MPC methods is alone a major factor in hindering their usage in practice (by demanding more-advanced processing units or inability to handle plants with fast dynamics). But on a positive note, as a by product of control law computations, an educated guess about the possible future behavior of dynamics with its corresponding input trajectory is available. This extra information can be viewed as a basis to design a triggering mechanism. It is worth mentioning that the idea of exploiting the computed optimal trajectories (from a computational savings perspective) is not entirely new in the MPC literature. The basic idea, the so-called *warm start*, is to properly shift the optimal input and state trajectories calculated in the previous instant. Then, one considers these shifted trajectories as the initial guess at the current instant to solve the optimization problem. By doing so, one can speed up the process of finding the optimal decision variables [54]. On a practical note, there is also a big incentive to exploit the computed optimal trajectories. Particularly in *wireless sensor/actuator networks* (WSANs), the single most important concern is the energy efficiency. Roughly speaking, the energy depletion source in a wireless node (e.g., placed on a sensor, actuator, or controller) is either from the micro-controller (responsible for logical/mathematical computations) or from the transceiver (responsible for sending and receiving data). The main source of energy depletion is however the transceiver. This observation is a key motivation to aggregate the data into a single packet (if possible) and transmit only the resulting packet over the communication network [55] and [56]. Thus, it becomes clear that an event-based policy to implement an MPC approach is naturally compatible with this nature of WSANs. For a detailed discussion on this subject, we refer the reader to [57] and the references therein (see [57, Subsection IV.B] for the energy efficiency topic). According to the above discussion, the combination of MPC methods with event-based implementation policies is in fact a (relevant) venue to explore.

Statement of contribution: In this chapter, an event-based sampling of an RMPC method is proposed for perturbed linear time-invariant (LTI) systems. The RMPC method is inspired by a constraint-tightening RMPC approach introduced in [58]. One key feature of the event-based method is its applicability on systems with decentralized sensing units. By the term *decentralized sensing*, we refer to systems where the sensory units are spatially dispersed. Although the sensory components' information can be collected on a single *centralized* node to place a triggering mechanism, we make the general assumptions that (i) the extra burden on the communication component is undesirable, and (ii) the complications that arise from the non-collocation of the triggering mechanism with the sensory components reduce the practicality of the triggering mechanism. Notice that these restrictions commonly occur in WSANs. Another important feature of our event-based implementation policy is the fact that the triggering mechanism's design is

decoupled from the design of the underlying RMPC method. In fact, the triggering mechanism is built upon the intrinsic robustness of the RMPC method. The core idea behind our method is to construct a sequence of *hyper-rectangles* around the optimal trajectories available from solving the RMPC problem (notice that these hyper-rectangles can be viewed as a certain type of *weighted* ℓ_∞ norms, see Definition 3.2.7). Then, these hyper-rectangles will be sent to the sensors and the optimal input trajectory will also be transmitted to the actuators. As soon as the observed states at the sensory units leave these hyper-rectangles, a triggering happens and the states at the triggering instant will be transmitted to the controller. Then, the process is repeated in a sampled-data fashion. The decentralized nature of the triggering mechanism stems from the fact that the Euclidean coordinate axes are normal (or perpendicular) to the faces of the constructed hyper-rectangles. Thus, the event-triggering mechanism can decide based on the local information, available at each individual sensor, whether to trigger or not. The main contributions of this chapter are summarized as follows.

- **Decoupled recursive feasibility and stability:** Given an RMPC method in place, we propose a set-theory-based triggering approach that preserves robust recursive feasibility and robust stability. Unlike the existing literature, the proposed approach is decoupled from the control synthesis process and does not require additional assumptions, such as extra conditions on eigenvalues of weighting matrices in the cost function or the need to define user-specified thresholds for the triggering mechanism (Theorems 3.4.3 & 3.4.4).
- **Decentralized nature:** The proposed approach enjoys a decentralized triggering mechanism that only requires local sensory information (Algorithm 2).
- **Tractable convex program reformulation:** We show that a certain type of non-convex volume-maximization problem with set-based constraints that is deployed to design the triggering mechanism admits a finite tractable convex program (CP) reformulation (Theorem 3.4.6).
- **Suboptimal linear program relaxation:** Motivated by an approach in the literature, we further show that a linear program (LP) relaxation of the CP reformulation is possible (Theorem 3.4.7).

Literature review: In what follows, we first review several approaches that are closely related to the problem considered in this chapter. We then give an account of algorithmic-oriented approaches in the literature that reduce the computational complexity of their MPC methods with customized algorithms.

Related works: In order to avoid repetition of terms (unless the contrary is mentioned otherwise), let us first mention the shared properties of the references below: linear discrete-time models, event-triggering mechanisms, constrained MPC methods, minimal (to none) coupling of the parameters of the triggering mechanism and the considered MPC method, and a computationally viable approach to design the triggering mechanism. Let us now elaborate the reasoning behind the last two properties. First, in an *ideal* case, one seeks the possibility of a complete *decoupling* between the parameters of triggering mechanism and of the considered MPC method. (By doing so, a fair comparison between the performances of the event-based and standard implementations

of an MPC method will be possible.) After all, our main goal is to provide an implementation policy for an MPC method with the awareness of communication issues. Hence, we only mention the approaches that have minimal to no inter-connections between the two sets of design parameters. Second, we also do not consider event-triggering approaches that are computationally more expensive compared to the underlying MPC problem (e.g., approaches that require to solve some type of an integer program). As mentioned above, the prohibitive computational requirement of MPC methods is the main factor that limits their application in practice. Thus, the viability of such complicated event-triggering designs becomes questionable in practice. The review of related works follows.

To deal with practical issues such as a band-limited communication channel, a novel design approach for NCSs is proposed in [59]. They employ the notion of *moving horizon* [60] to design the estimator and controller. A remarkable character of their approach is its ability to decide *on-the-fly* which input channel should be updated (i.e., a certain type *input-channel* event-triggering control). In case of collocated controller and actuator units, an event-based estimator with a bounded covariance matrix is designed in [61]. While the estimator receives data via a Lebesgue sampling approach, it periodically updates the controller's information regarding the disturbances with a polytopic overapproximation of covariance matrix. The authors of [62] propose an interesting transmission strategy for wireless sensor/controller communications with practical energy-aware provisions (the controller is collocated with the actuator system). Using some predefined thresholds for each state's sensor (i.e., an ℓ_1-type triggering mechanism), the controller is computed offline using an *explicit* MPC approach [63]. Based on a prescribed 2-norm ball around the optimal state trajectory, the authors in [64] propose a triggering mechanism for WSANs. They show that the approach is robustly stable to a set that is a function of the radius of threshold ball and the maximal 2-norm of disturbance. For linear, continuous-time dynamical systems affected by a Wiener process, a co-design method (i.e., simultaneous design of the scheduler and the controller) is proposed in [65]. The main idea is inspired by the notion of *rollout* from *dynamic programming* [66]. More importantly, the authors show that under some mild conditions an event-based control approach outperforms a traditional control approach in the sense of closed-loop performance/average transmission rate. (Notice that for most of the approaches in the literature including this chapter such a guarantee is not provided.) A set theoretic triggering mechanism is introduced in [67] for systems with collocated controller and sensory units. The approach is inspired by the *tube-based* MPC proposed in [68]. By exploiting the known probability distribution of disturbance, they also guarantee an average sampling rate. However, their tube-contraction method requires a certain type of realization of a discrete-time system, see [67, Remark 8]. Demirel et al., introduce a sensor/actuator event-triggering mechanism for control systems with limited number of control messages (i.e., communication and computation resources are scarce) [69]. They relax the original combinatorial problem into a convex one by an appropriate definition of event thresholds.

Algorithmic viewpoint: The reasoning behind this algorithmic viewpoint is as follows. An MPC optimization problem is computationally expensive by itself (the evidence is the substantial body of work that has been done to customize algorithms to MPC prob-

lems). Hence, the merit of an event-based policy of implementation would be lost if the mechanism demands a drastically higher computational effort compared to the underlying MPC problem. We should highlight that all the approaches discussed next are related to linear time-invariant (LTI) systems. Dunn and Bertsekas in [70] exploit the structure of their considered optimization problem to reduce the cost of Newton's step in their dynamic programming approach. They show that the arithmetic cost of a Newton step in their approach scales as a linear function of the horizon length N instead of as a cubic function of it. In [71], the authors use the *primal barrier interior-point method* with a specific type of *ordering of decision variables*. By doing so, they show that the underlying problem possesses a desirable sparse structure that decreases the computational cost significantly to compute the control action. The authors in [72] employ the celebrated *Nesterov's accelerated method* to solve the underlying optimization problem. Inherited from Nesterov's method (that is a gradient-based algorithm), their approach possesses a simple implementation requirement. More importantly, they also provide *a priori* computational complexity certificate. There is also another class of MPC approaches in the literature that can be computed *offline* and employed for systems with fast dynamics, the so-called explicit MPC methods [63]. Here, the solution of the MPC problem is constructed as a function of the initial state in a form of a *lookup table*. This lookup table is then utilized in an online fashion. However, since the size of this table is exponentially dependent on the decision variables' dimension and the horizon length, this method's applicability is limited to low dimensions.

The layout of the chapter is as follows. The mathematical notions used in the chapter are outlined in Section 3.2. Section 3.3 is devoted to the considered RMPC method. The main results regarding the event-based implementation policy are introduced in Section 3.4. Section 3.5 contains the technical proofs. A numerical example is presented in Section 3.6 to evaluate the effectiveness of theoretical results. Finally, the chapter is concluded in Section 3.7.

Notation: the set of non-negative integers is denoted by $\mathbb{Z}_{\geq 0}$. Given positive integers m and n, $\mathbb{R}^m$ and $\mathbb{R}^{m\times n}$ represent the m-dimensional Euclidean space and the space of $m \times n$ matrices with real entries, respectively. Given a positive integer r, the sets of positive integers and non-negative integers less than or equal to r are denoted by $\mathsf{N}_{[r]}$ and $\mathsf{Z}_{[r]}$, respectively. Given a vector $v \in \mathbb{R}^n$, v^i represents the i-th entry of v. For any pairs of vectors $a, b \in \mathbb{R}^n$, the inequality $a < (\leq) b$ is realized in a componentwise manner, i.e., $a^i < (\leq) b^i$, for all $i \in \mathsf{N}_{[n]}$. Given a matrix $M \in \mathbb{R}^{m\times n}$, M_{ij} denotes the i-th row, j-th column entry of M. Moreover, the matrix $M^+ \in \mathbb{R}^{m\times n}$ is the matrix with entries of $M^+_{ij} := \max\{0, M_{ij}\}$. A positive definite matrix M is denoted by $M \succ 0$. The $n \times n$ identity matrix is denoted by I_n. Given a vector $v \in \mathbb{R}^n$ and a scalar $p \geq 1$, $\|v\|_p$ denotes the p-norm $\left(\sum_{i=1}^n (v^i)^p\right)^{1/p}$. The function $\mathrm{sign}(\cdot)$ represents the standard sign function. Given a set $\mathscr{S} \subset \mathbb{R}^n$ and a matrix $M \in \mathbb{R}^{m\times n}$, the set $M\mathscr{S}$ denotes the set $\{c \in \mathbb{R}^m : \text{ there exist } s \in \mathscr{S} \text{ such that } Ms = c\}$. Given two sets $\mathscr{A}$ and $\mathscr{B}$ in $\mathbb{R}^n$, $\mathscr{A} \,/\, \mathscr{B} := \{x \in \mathscr{A} : x \notin \mathscr{B}\}$.

3.2. PRELIMINARIES

In what follows, we begin with a brief review of the mathematical preliminaries used in the rest of the chapter. The section is divided to two parts: RMPC related notions and Set theory related notions.

RMPC related notions: We first review some notions that the RMPC method is based on.

Definition 3.2.1 (Point-to-set weighted distance)**.** *Given a matrix $M \succ 0$, the squared weighted distance of a point $r \in \mathbb{R}^n$ from a set $\mathscr{S} \subset \mathbb{R}^n$ is defined as*

$$d(r, \mathscr{S}, M) := \min_{s \in \mathscr{S}} ||r - s||_M^2 = \min_{s \in \mathscr{S}} (r-s)^\top M (r-s).$$

When $\mathscr{S}$ is singular, i.e., $\mathscr{S} = \{s\}$, the distance is

$$d(r, \{s\}, M) = ||r - s||_M^2 = (r-s)^\top M (r-s).$$

Definition 3.2.2 (Pontryagin difference and Minkowski sum)**.** *Given sets $\mathscr{C}$ and $\mathscr{D}$, the Pontryagin difference $\mathscr{C} \sim \mathscr{D}$ and the Minkowski sum $\mathscr{C} \oplus \mathscr{D}$ are defined as*

$$\mathscr{C} \sim \mathscr{D} := \{c : \ c + d \in \mathscr{C}, \text{ for all } d \in \mathscr{D}\},$$
$$\mathscr{C} \oplus \mathscr{D} := \{c + d : \text{ for all } c \in \mathscr{C}, \text{ for all } d \in \mathscr{D}\}.$$

The following result will be used extensively in the development of the triggering mechanism.

Lemma 3.2.3 ([58])**.** *Let a be a vector in $\mathbb{R}^n$, $\mathscr{B}$ and $\mathscr{C}$ be two compact sets in $\mathbb{R}^n$, and M be a positive definite matrix in $\mathbb{R}^{n \times n}$. Then, using the distance function given in Definition 3.2.1, we have*

$$d(a + c, \mathscr{B}, M) \leq d(a, \mathscr{B} \sim \mathscr{C}, M), \quad \text{for all } c \in \mathscr{C}.$$

We next introduce a procedure to compute a certain type of gains that play an essential role in the constraint-tightening approach proposed in [58].

Definition 3.2.4 (M-Step nilpotent LQR controller [58])**.** *Given two positive definite matrices Q and R, and two positive integers M and N such that $M \leq N-1$. Consider the controllable LTI system $\xi_{k+1} = A\xi_k + B\nu_k$, where $\xi \in \mathbb{R}^n$ and $\nu \in \mathbb{R}^m$. The following backward recursion produces as output a set of linear state feedback gains $\mathbf{K} = \{K_i\}_{i=0}^{N-1}$ that drives the state of the nominal system (3.4) to the origin in M steps, and remains there until step N.*

1. *Set $K_j = 0_{m \times n}$, for all $j \in \{M, \ldots, N-1\}$.*

2. *Set $P_M = 0_{n \times n}$, $S_M = \mathsf{I}_n$.*

3. Compute in backward for $j \in \{M-1, M-2, \ldots, 0\}$,

$$
\begin{aligned}
K_j &= -\begin{pmatrix} \mathsf{I}_m & 0_{m\times n}\end{pmatrix} H_{j+1}^{\dagger} \begin{pmatrix} B^\top P_{j+1} \\ S_{j+1}\end{pmatrix} A, \\
H_{j+1} &= \begin{pmatrix} (R + (B^\top P_{j+1} B)) & B^\top S_{j+1} \\ S_{j+1} B & 0_{n\times n}\end{pmatrix}, \\
S_j &= (A + BK_j)^\top S_{j+1} (A + BK_j), \\
P_j &= Q + K_j^\top R K_j + (A + BK_j)^\top P_{j+1} (A + BK_j).
\end{aligned}
$$

Set theory related notions: We now recall some notions from convex analysis (see e.g., [73, Section 2] for a compact exposition of the subject).

Definition 3.2.5 (Support function)**.** *Given a set $\mathscr{S} \subset \mathbb{R}^n$, the support function of $\mathscr{S}$ evaluated at $\eta \in \mathbb{R}^n$ is*

$$h_{\mathscr{S}}(\eta) = \sup_{s\in\mathscr{S}} \langle \eta, s\rangle.$$

The domain $\mathscr{K}_{\mathscr{S}}$ on which the support function is defined is a convex cone pointed at the origin. If $\mathscr{S}$ is bounded, then $\mathscr{K}_{\mathscr{S}} = \mathbb{R}^n$. Given a matrix $M \in \mathbb{R}^{n\times m}$ and a vector $v \in \mathbb{R}^n$ such that $G^\top v \in \mathscr{K}_{\mathscr{S}}$,

$$h_{M\mathscr{S}}(v) := h_{\mathscr{S}}(M^\top v).$$

Suppose $\mathscr{S} \subset \mathbb{R}^n$ is closed and convex. Then,

$$\mathscr{S} := \{s \in \mathbb{R}^n : \langle \eta, s\rangle \le h_{\mathscr{S}}(\eta), \text{ for all } \eta \in \mathscr{K}_{\mathscr{S}}\},$$

i.e., the intersection of its supporting half spaces.

Definition 3.2.6 (Polyhedron)**.** *A set $\mathscr{S} \subset \mathbb{R}^n$ is called a polyhedron, if*

$$\mathscr{S} := \{s \in \mathbb{R}^n : A_{\mathscr{S}} s \le b_{\mathscr{S}}\},\ A_{\mathscr{S}} \in \mathbb{R}^{m\times n},\ b_{\mathscr{S}} \in \mathbb{R}^m.$$

If the polyhedron $\mathscr{S}$ is bounded, the set is also called a *polytope* and the representation given in Definition 3.2.6 is known as the *H-representation*. Furthermore, the support function $h_{\mathscr{S}}(\eta)$ is the solution of the LP,

$$
\begin{aligned}
h_{\mathscr{S}}(\eta) = \max_{s}\ & \langle \eta, s\rangle \\
\text{s.t. } & A_{\mathscr{S}} s \le b_{\mathscr{S}}.
\end{aligned}
$$

Given the H-representation of a polytope, we employ the notations $a_{i,\mathscr{S}} \in \mathbb{R}^{1\times n}$ and $a_{\mathscr{S},j} \in \mathbb{R}^{m\times 1}$ to denote the i-th row and the j-th column of $A_{\mathscr{S}}$, respectively.

Consider the polyhedron $\mathscr{S}$ given by Definition 3.2.6 and a set $\mathscr{V} \subset \mathbb{R}^n$. Assuming $h_{\mathscr{V}}(a_{i,\mathscr{S}}^\top)$ is well-defined for all $i \in \mathsf{N}_{[m]}$. Then,

$$\mathscr{S} \sim \mathscr{V} := \{z \in \mathbb{R}^n : \langle a_{i,\mathscr{S}}^\top, z\rangle \le b_{i,\mathscr{S}} - h_{\mathscr{V}}(a_{i,\mathscr{S}}^\top), \text{ for all } i \in \mathsf{N}_{[m]}\}, \tag{3.1}$$

where $a_{i,\mathscr{S}}$ and $b_{i,\mathscr{S}}$ are the i-th row of $A_{\mathscr{S}}$ and the i-th entry of $b_{\mathscr{S}}$, respectively.

Definition 3.2.7 (Full-dimensional hyper-rectangle)**.** *For any vector-pairs $l, u \in \mathbb{R}^n$ such that $l < u$, the full-dimensional convex polytope*

$$\begin{aligned}\mathscr{B}(l,u) &:= \{x \in \mathbb{R}^n : \ l \le x \le u\} \\ &= \{x \in \mathbb{R}^n : \ A_{\mathscr{B}} x \le b_{\mathscr{B}}\}\end{aligned}$$

is called a hyper-rectangle*, where $A_{\mathscr{B}} := [\mathsf{I}_n \ -\mathsf{I}_n]^\top$ and $b_{\mathscr{B}} = [u^\top \ -l^\top]^\top$.*

3.3. RMPC METHOD

In this section, we introduce the class of constrained dynamical systems considered in this chapter, followed by the description of the RMPC method. The description contains the main ingredients of the RMPC method which play crucial roles in the design of the decentralized event-triggering mechanism. At last, we formally state the problem addressed in this chapter.

3.3.1. SYSTEM DESCRIPTION

Consider an LTI system with bounded additive perturbations given by

$$x_{k+1} = Ax_k + Bu_k + w_k, \text{ for all } k \in \mathbb{Z}_{\ge 0}, \tag{3.2}$$

where the state, input and disturbance signals satisfy the (*hard*) constraints

$$x_k \in \mathbb{X} \subset \mathbb{R}^{n_x}, \ u_k \in \mathbb{U} \subset \mathbb{R}^{n_u}, \ w_k \in \mathscr{W} \subset \mathbb{R}^{n_x}. \tag{3.3}$$

The nominal system associated with (3.2) is

$$\bar{x}_{k+1} = A\bar{x}_k + Bu_k, \text{ for all } k \in \mathbb{Z}_{\ge 0}. \tag{3.4}$$

The RMPC method is designed such that the state x_k and the input u_k converge to some *target sets* $\mathbb{T}_x \subset \mathbb{R}^{n_x}$ and $\mathbb{T}_u \subset \mathbb{R}^{n_u}$ as $k \to \infty$, respectively, while the constraints (3.3) are satisfied at all instants.

Assumption 3.3.1 (System & set properties)**.** *Consider the dynamics* (3.2) *with the constraints* (3.3)*. We suppose that the following conditions hold.*

- *(Nominal controllablity) The pair (A, B) is controllable.*
- *(Polytopic sets) The constraint sets $\mathbb{X}$ and $\mathbb{U}$, the target sets $\mathbb{T}_x$ and $\mathbb{T}_u$, and the disturbance set $\mathscr{W}$ are all convex, compact polytopes containing their underlying spaces' origin in their interior.*

3.3.2. RMPC FORMULATION

We start with introducing two types of feedback gains which are essential to the RMPC method and the construction of the triggering mechanism, as well.

1) Nominal feedback gain: The first item in Assumption 3.3.1 implies that the discrete algebraic Riccati equation [74]

$$P = A^\top PA - (A^\top PB)(R + B^\top PB)^{-1}(B^\top PA) + Q,$$

associated with the nominal system (3.4) attains a unique solution $P \succ 0$, where the positive definite matrices Q and R are properly chosen (i.e., the pair $(A, Q^{\frac{1}{2}})$ is detectable). One can then employ P to find a stabilizing state-feedback gain

$$F = -(R + B^\top P B)^{-1} B^\top A. \tag{3.5}$$

2) Disturbance feedback gains: Let integer N be the horizon length of the RMPC method and positive integer M be given, where $N \geq n_x + 1$ and $n_x \leq M \leq N-1$. In the RMPC method, the tightening approach is based on a set of disturbance gains $\mathbf{K} = \{K_i\}_{i=0}^{N-1}$. We employ the procedure introduced in Definition 3.2.4 to construct the set $\mathbf{K}$.

Remark 3.3.2 (Roles of feedback gains)**.** *The state-feedback gain F is related to the states that lie inside the terminal set of the RMPC method. This gain plays an active role in the proofs of robust recursive feasibility and robust stability (see the second ingredient of the RMPC method introduced below). On the other hand, there exists a model mismatch between the perturbed system* (3.2) *and the nominal system* (3.4)*. Thus, the constructed state trajectories using the nominal system* (3.4) *suffer from prediction errors in the RMPC problem. By means of these gains, the RMPC method takes into account the effect of unknown disturbances. As a result, the method can guarantee the predicted trajectories satisfy the constraints of the RMPC problem (see the first ingredient of the RMPC method introduced below).*

Ingredients of RMPC: We now present concise descriptions of three involved ingredients of the RMPC method: the constraint tightening mechanism, the terminal set, and the cost function. Suppose $k \in \mathbb{Z}_{\geq 0}$ is the instant at which the RMPC problem, denoted by $\mathscr{P}(x_k)$, is solved. Let the positive integer N be the length of the finite horizon. Consider the nominal feedback gain F given in (3.5) and the disturbance gains $\mathbf{K}$ computed based on Definition 3.2.4, both for the nominal dynamics (3.4).

(1) Constraint tightening mechanism: The following rule of constraint tightening is applied to the input, state, input target, and state target sets, for all $i \in \mathsf{Z}_{[N-2]}$,

$$\mathscr{U}_0 = \mathbb{U}, \qquad \mathscr{U}_{i+1} = \mathscr{U}_i \sim K_i L_i \mathscr{W}, \tag{3.6a}$$
$$\mathscr{X}_0 = \mathbb{X}, \qquad \mathscr{X}_{i+1} = \mathscr{X}_i \sim L_i \mathscr{W}, \tag{3.6b}$$
$$\mathscr{T}_{u,0} = \mathbb{T}_u, \quad \mathscr{T}_{u,i+1} = \mathscr{T}_{u,i} \sim K_i L_i \mathscr{W}, \tag{3.6c}$$
$$\mathscr{T}_{x,0} = \mathbb{T}_x, \quad \mathscr{T}_{x,i+1} = \mathscr{T}_{x,i} \sim L_i \mathscr{W}, \tag{3.6d}$$

where

$$L_0 = \mathsf{I}_{n_x}, \quad L_{i+1} = (A + BK_i) L_i. \tag{3.6e}$$

(2) Terminal set: Let $\mathscr{R}$ be a robust control invariant set under the disturbances

$L_{N-1}\mathscr{W}$ such that if $x \in \mathscr{R}$ [1],

$$(A+BF)x + L_{N-1}w \in \mathscr{R}, \text{ for all } w \in \mathscr{W}, \tag{3.7a}$$
$$x \in \mathscr{X}_{N-1}, \tag{3.7b}$$
$$x \in \mathscr{T}_{x,N-1}, \tag{3.7c}$$
$$Fx \in \mathscr{U}_{N-1}, \tag{3.7d}$$
$$Fx \in \mathscr{T}_{u,N-1}. \tag{3.7e}$$

The terminal state set is then

$$\mathscr{X}_f = \mathscr{R} \sim L_{N-1}\mathscr{W} \subset \mathbb{R}^{n_x}. \tag{3.8}$$

Assumption 3.3.3. *The terminal set $\mathscr{X}_f$ is non-empty.*

(3) Cost function: Let us denote the input trajectory $\{u_{k+i|k}\}_{i=0}^{N-1}$ and the state trajectory $\{x_{k+i|k}\}_{i=0}^{N}$ by $\mathbf{U}_{k|k}$ and $\mathbf{X}_{k|k}$, respectively. The cost function of the RMPC problem is

$$J(x_k, \mathbf{U}_{k|k}) = \sum_{i=0}^{N-1} d(\bar{x}_{k+i|k}, \mathscr{T}_{x,i}, Q) + d(u_{k+i|k}, \mathscr{T}_{u,i}, R), \tag{3.9}$$

where d is the weighted distance function introduced in Definition 3.2.1.

We are now in a position to introduce the RMPC problem.

RMPC formulation: The optimization problem $\mathscr{P}(x_k)$ for a finite horizon N at the instant k reads as follows:

$$\mathscr{P}(x_k) : J(x_k, \mathbf{U}^*_{k|k}) = \min_{\mathbf{X}_{k|k}, \mathbf{U}_{k|k}} J(x_k, \mathbf{U}_{k|k}) \tag{3.10a}$$

subject to

$$\bar{x}_k = x_k \tag{3.10b}$$
$$\bar{x}_{k+i+1|k} = A\bar{x}_{k+i|k} + Bu_{k+i|k}, \text{ for all } i \in \mathsf{Z}_{[N-1]} \tag{3.10c}$$
$$u_{k+i|k} \in \mathscr{U}_i, \text{ for all } i \in \mathsf{Z}_{[N-1]} \tag{3.10d}$$
$$\bar{x}_{k+i|k} \in \mathscr{X}_i, \text{ for all } i \in \mathsf{Z}_{[N-1]} \tag{3.10e}$$
$$\bar{x}_{k+N|k} \in \mathscr{X}_f, \tag{3.10f}$$

where $\mathbf{U}^*_{k|k} := \{u^*_{k+i|k}\}_{i=0}^{N-1}$ and $\mathbf{X}^*_{k|k} := \{x^*_{k+i|k}\}_{i=0}^{N}$ represent the optimal input and state trajectories of $\mathscr{P}(x_k)$, respectively.

In what follows, we state the preliminary results of this chapter concerning robust recursive feasibility and robust convergence of the RMPC method.

Theorem 3.3.4 (Robust Recursive Feasibility)**.** *Consider the perturbed LTI dynamics* (3.2) *subject to the constraints and the disturbances given in* (3.3)*. Suppose that for some initial state $x_0 \in \mathbb{R}^{n_x}$, $\mathscr{P}(x_0)$ has a feasible solution. Then, (i) the following optimization problems $\mathscr{P}(x_k)$ have feasible solutions for all $k \in \mathbb{Z}_{\geq 0}$. Additionally, (ii) the trajectories of the system (3.2) satisfy the constraints in (3.3).*

[1] Notice that when the disturbance feedback gains $\mathbf{K}$ render the system *nilpotent* in less than $N-1$ steps, it follows that $L_{N-1} = 0$. Hence, the conditions on the terminal set $\mathscr{X}_f$ become less restrictive, i.e., see (3.8) in this regard.

Theorem 3.3.5 (Robust convergence)**.** *Consider the perturbed LTI dynamics* (3.2) *subject to the constraints and the disturbances given in* (3.3)*. Suppose that for some initial state* $x_0 \in \mathbb{R}^{n_x}$, $\mathscr{P}(x_0)$ *has a feasible solution. Then, the state and input trajectories are such that* $x_k \to \mathbb{T}_x$ *and* $u_k \to \mathbb{T}_u$*, as* $k \to \infty$*.*

The proofs of Theorems 3.3.4 and 3.3.5 are provided in Subsections 3.5.1 and 3.5.2, respectively. The main problem addressed in this chapter is now introduced.

Problem 3.3.6 (Formal statement)**.** *Consider the perturbed LTI dynamics* (3.2) *with the constraints* (3.3)*. Suppose now the sensory units are decentralized (in the sense that there is no centralized unit outside the controller unit that collects all the states). Under Assumptions 3.3.1 and 3.3.3, the controller unit synthesizes an optimal input trajectory* $\mathbf{U}^*_{k|k}$*, where* $k \in \mathbb{Z}_{\geq 0}$ *is the last instant at which the optimization problem* (3.10) *is solved.*

Devise a triggering mechanism to determine the next triggering instant $k_{\text{trig}} > k$ *in the form of*

$$k_{\text{trig}} := k + \min\{j \in \mathsf{N}_{[N-1]} : \; x_{x+j} \notin x^*_{k+j|k} \oplus \mathscr{E}_{j,k}\}, \tag{3.11}$$

such that

- *the actuator units employ the control action* $u_{k+j} = u^*_{k+j|k}$*, for all* $j \in \mathsf{Z}_{[k_{\text{trig}}-k]}$*;*
- *the closed-loop dynamics* (3.2) *with the above control law does not violate the constraints* (3.3)*;*
- *the states* $x_k \to \mathbb{T}_x$ *and the inputs* $u_k \to \mathbb{T}_u$*, as* $k \to \infty$*,*

where x_{k+j} *is the observed state at the sensory units and the sequence of hyper-rectangle sets* $\mathscr{E}_k := \{\mathscr{E}_{j,k}\}_{j=1}^{N-1}$ *ought to be determined.*

Remark 3.3.7 (Standard vs event-based implementation policies)**.** *In a standard RMPC setting, such as the problem* (3.10)*, the optimal control problem is solved at all instants* $k \in \mathbb{Z}_{\geq 0}$*. The first element of* $\mathbf{U}^*_{k|k}$*, that is* $u^*_{k|k}$*, is then applied to the plant via the actuator nodes at every instant* k*. While in an event-based setting, the triggering mechanism generally exploits the optimal trajectories* $\mathbf{U}^*_{k|k}$ *and* $\mathbf{X}^*_{k|k}$ *to provide sufficient conditions under which robust stability and robust recursive feasibility are satisfied.*

Remark 3.3.8 (Iteration Complexity of RMPC)**.** *An RMPC problem with linear dynamics, a quadratic cost function, and polytopic constraints is a* quadratic program *(QP) (the cost function is quadratic in decision variables and constraints are given by linear equalities or inequalities in terms of decision variables). Interestingly, dedicated solvers can provide the complexity per iteration* $\mathscr{O}(N(n_x + n_u)^3)$ *by exploiting the structure of the corresponding QP [71]. Notice that a common approach to solve QPs is the interior-point method [75] which can guarantee the complexity per iteration* $\mathscr{O}(N^3(n_x + n_u)^3)$*. It is however worth noting that because of the type of the cost function* (3.9)*, the resulting QP has the complexity per iteration* $\mathscr{O}(N(2n_x + 2n_u)^3)$*.*

3.4. MAIN RESULTS

We provide the event-based implementation policy of the RMPC method (3.10) in this section. We first describe a certain type of constrained optimization problems. The outcome of these optimization problems is a sequence of hyper-rectangles $\mathscr{E}_k := \{\mathscr{E}_{j,k}\}_{j=1}^{N-1}$. Based on these constructed sets, we then state the main results of this chapter.

3.4.1. CONSTRUCTION OF HYPER-RECTANGLES $\mathscr{E}_\mathbf{k}$

We introduce the framework to construct each hyper-rectangle $\mathscr{E}_{j,k}$ which includes: the chosen type of parameterization of each $\mathscr{E}_{j,k}$, the required quantities to cast the problem of finding $\mathscr{E}_{j,k}$ as an optimization problem, followed by the optimization problem itself. Suppose k is the last instant at which the problem $\mathscr{P}(x_k)$ has been solved. Let $j \in \mathsf{N}_{[N-1]}$ denote a time instant following k. Notice that j indicates the instant at which the triggering mechanism is being evaluated. Define the prediction error

$$e_{k+j|k} = x_{k+j} - x^*_{k+j|k}, \tag{3.12}$$

due to the mismatch between the perturbed dynamics and the nominal dynamics. For all $p \in \mathsf{N}_{[n_x]}$, x^p_{k+j} is the p-th element of the state vector x_{k+j}, observed by the p-th local sensor. For all $j \in \mathsf{N}_{[N-1]}$ and $p \in \mathsf{N}_{[n_x]}$, consider some positive scalars $\underline{e}^p_{j,k}$ and $\overline{e}^p_{j,k}$ defining $(N-1)$ hyper-rectangles

$$\mathscr{E}_{j,k} := \left\{\epsilon \in \mathbb{R}^{n_x} : -\underline{e}^p_{j,k} \le \epsilon^p \le \bar{e}^p_{j,k}, \text{ for all } p \in \mathsf{N}_{[n_x]}\right\} \subset \mathbb{R}^{n_x},$$

where ϵ^p is the p-th element of ϵ. In simple words, each hyper-rectangle $\mathscr{E}_{j,k}$ is parameterized in $2n_x$ parameters $\underline{e}^p_{j,k}$ and $\overline{e}^p_{j,k}$. In the next step, we introduce a set-based optimization problem to find each set $\mathscr{E}_{j,k}$ such that if $e_{k+j|k} \in \mathscr{E}_{j,k}$, then, the three requirements in Problem 3.3.6 are fulfilled. Based on the error definition (3.12), one can reformulate the triggering mechanism (3.11) as $k_{\text{trig}} := k + \min\{j \in \mathsf{N}_{[N-1]} : e_{k+j|k} \notin \mathscr{E}_{j,k}\}$.

First, let us introduce the quantities involved in the derivation of the sets $\mathscr{E}_{j,k}$. Recall that $\mathbf{U}^*_{k|k} = \{u^*_{k+i|k}\}_{i=0}^{N-1}$ and $\mathbf{X}^*_{k|k} = \{x^*_{k+i|k}\}_{i=0}^{N}$ are available from solving $\mathscr{P}(x_k)$.

(1) Construction of optimal input and state trajectories: We now aim to construct the following (with some abuse of notation) optimal trajectories $\mathbf{U}^*_{k+j|k} := \{u^*_{k+j+i|k}\}_{i=0}^{N-1}$, and $\mathbf{X}^*_{k+j|k} := \{x^*_{k+j+i|k}\}_{i=0}^{N}$. The term abuse of notation refers to the fact that we have only access to $u^*_{k+j+i|k}$ for $j+i \le N-1$ and $x^*_{k+j+i|k}$ for $j+i \le N$ from solving $\mathscr{P}(x_k)$. We adopt the convention

$$u^*_{k+j+i|k} = F(A+BF)^{j+i-N} x^*_{k+N|k}, \qquad \text{for } j+i \ge N, \tag{3.13a}$$

$$x^*_{k+j+i|k} = (A+BF)^{j+i-N} x^*_{k+N|k}, \qquad \text{for } j+i \ge N+1, \tag{3.13b}$$

to construct unavailable "optimal" input and state trajectories based on the solution of $\mathscr{P}(x_k)$.

(2) Construction of minimal distance target points: Let us now introduce two new

sets of quantities

$$s^*_{x,k+j+i|k} = \underset{s_x \in \mathcal{T}_{x,j+i}}{\operatorname{argmin}} \| x^*_{k+j+i|k} - s_x \|^2_Q, \text{ for } j+i \leq N, \tag{3.14a}$$

$$s^*_{u,k+j+i|k} = \underset{s_u \in \mathcal{T}_{u,j+i}}{\operatorname{argmin}} \| u^*_{k+j+i|k} - s_u \|^2_R, \text{ for } j+i \leq N-1. \tag{3.14b}$$

Observe that the points $s^*_{x,k+j+i|k}$ and $s^*_{u,k+j+i|k}$ represent the points on the boundary of the target sets that have the smallest distance to $x^*_{k+j+i|k}$ and $u^*_{k+j+i|k}$, respectively. These parameters are already at hand as a by product of solving $\mathscr{P}(x_k)$. We next define

$$s^*_{x,k+j+i|k} = x^*_{k+j+i|k}, \text{ for } j+i \in \mathsf{Z}_{[N+j]}/\mathsf{Z}_{[N]} \tag{3.15a}$$

$$s^*_{u,k+j+i|k} = u^*_{k+j+i|k}, \text{ for } j+i \in \mathsf{Z}_{[N+j-1]}/\mathsf{Z}_{[N-1]}, \tag{3.15b}$$

for all $i \in \mathsf{Z}_{[N-1]}$. The convention introduced in (3.13) implies that $u^*_{k+j+i|k} \in \mathcal{T}_{u,i}$ and $x^*_{k+j+i|k} \in \mathcal{T}_{x,i}$, for $j+i > N-1$ (since they are state-feedback extensions of the terminal state $x^*_{k+N|k}$). In light of this fact, $d(x^*_{k+j+i|k}, \mathcal{T}_{x,i}, Q) = d(u^*_{k+j+i|k}, \mathcal{T}_{u,i}, R) = 0$ and hence, the choice made in the definition (3.15) becomes apparent.

(3) Adopted feedback gains and transition matrices: We finally adopt the feedback gains $\tilde{K}_i$ and the state-transition matrices $\tilde{L}_i$ defined as[2]

$$\tilde{K}_0 = 0_{n_u \times n_x}, \qquad \tilde{K}_{i+1} = K_i, \qquad \text{for all } i \in \mathsf{Z}_{[N-2]}, \tag{3.16a}$$

$$\tilde{L}_0 = \mathsf{I}_{n_x}, \qquad \tilde{L}_{i+1} = (A + B\tilde{K}_i)\tilde{L}_i, \qquad \text{for all } i \in \mathsf{Z}_{[N-1]}. \tag{3.16b}$$

These gains and transition matrices enable us to construct candidate input and state trajectories at instant $k+j$ based on the optimal solution trajectories of $\mathscr{P}(x_k)$. We then utilize these candidate trajectories to prove the main results (see Subsections 3.5.3 and 3.5.4).

Construction of $\mathscr{E}_{j,k}$: Let us first provide two definitions for the volume of $\mathscr{E}_{j,k}$, that are

$$\text{vol}_1(\mathscr{E}_{j,k}) := \prod_{p \in [n_x]} (\overline{e}^p_{j,k} + \underline{e}^p_{j,k}), \tag{3.17a}$$

$$\text{vol}_2(\mathscr{E}_{j,k}) := \prod_{p \in [n_x]} (\overline{e}^p_{j,k} \times \underline{e}^p_{j,k}). \tag{3.17b}$$

Notice that the definition (3.17a) is the standard definition of the volume for $\mathscr{E}_{j,k}$ in $\mathbb{R}^{n_x}$. As it will be discussed later on, the constructed hyper-rectangles $\mathscr{E}_{j,k}$ based on $\text{vol}_1(\mathscr{E}_{j,k})$ are highly asymmetric with respect to the origin although these sets have the maximum possible volume. (This asymmetry in turn implies that the triggering mechanism has no robustness in certain error directions.) The non-standard definition (3.17b) of the volume is introduced to handle the asymmetry issue and to promote a more symmetric construction of $\mathscr{E}_{j,k}$ around the origin compared to the construction based on the definition (3.17a). The discussion regarding the importance of this point is provided in

[2] Notice that $\tilde{L}_1 = A$: the error evolves in open-loop for one time step.

Remark 3.4.9. The optimization problem to find the hyper-rectangle $\mathscr{E}_{j,k}$ follows:

$$\max_{\overline{e}^p_{j,k},\underline{e}^p_{j,k}} \ \mathrm{vol}_q(\mathscr{E}_{j,k}) \tag{3.18a}$$

subject to

$$\underline{e}^p_{j,k} \geq 0, \bar{e}^p_{j,k} \geq 0, \qquad \text{for all } p \in \mathsf{N}_{[n_x]}, \tag{3.18b}$$

$$x^*_{k+j+i|k} \in \mathscr{X}_i \sim \tilde{L}_i \mathscr{E}_{j,k}, \qquad \text{for all } i \in \mathsf{Z}_{[N-1]}, \tag{3.18c}$$

$$u^*_{k+j+i|k} \in \mathscr{U}_i \sim \tilde{K}_i \tilde{L}_i \mathscr{E}_{j,k}, \qquad \text{for all } i \in \mathsf{Z}_{[N-1]}, \tag{3.18d}$$

$$s^*_{x,k+j+i|k} \in \mathscr{T}_{x,i} \sim \tilde{L}_i \mathscr{E}_{j,k}, \qquad \text{for all } i \in \mathsf{Z}_{[N-1]}, \tag{3.18e}$$

$$s^*_{u,k+j+i|k} \in \mathscr{T}_{u,i} \sim \tilde{K}_i \tilde{L}_i \mathscr{E}_{j,k}, \qquad \text{for all } i \in \mathsf{Z}_{[N-1]}, \tag{3.18f}$$

where $q \in \{1,2\}$ determines which type of the volume definition in (3.17) is chosen. At the triggering instant k, all the required data to solve the problem (3.18) is available after solving the RMPC problem (3.10), for all $j \in \mathsf{N}_{[N-1]}$.

Remark 3.4.1 (Non-convexity and parametric-in-set constraints)**.** *The problem* (3.18) *simply seeks to find the maximum volume hyper-rectangles $\mathscr{E}_{j,k}$ that satisfy the set-based constraints* (3.18c)-(3.18f)*. It is evident that the objective function* $\mathrm{vol}_q(\mathscr{E}_{j,k})$ *is a nonlinear, non-convex function that makes the problem* (3.18) *difficult to solve by itself. Despite this non-convexity, we show that the problem remains practically solvable. We show that: (i) the set-based constraints* (3.18c)-(3.18f) *are effectively representable by linear inequalities (i.e., polytopic inequalities) in Theorems 3.4.6 & 3.4.7, (ii) the cost function* $\mathrm{vol}_q(\mathscr{E}_{j,k})$ *of the problem* (3.18) *has a CP counterpart in Theorem 3.4.6 , and (iii) the problem* (3.18) *admits an LP relaxation in Theorem 3.4.7.*

3.4.2. EVENT-BASED DECENTRALIZED IMPLEMENTATION

The theoretical results of this chapter are now introduced. We first present an algorithmic implementation of the RMPC method using the outcome of the optimization problem (3.18) (In Theorems 3.4.6 & 3.4.7, we will show how to solve the problem (3.18) efficiently). Then, we show that robust recursive feasibility and robust stability are guaranteed under the utilization of such an implementation policy. In simple words, if the incurred prediction errors caused by the model mismatch between two consecutive triggering instants are inside the hyper-rectangles $\mathscr{E}_k$, then, the decentralized, event-based implementation of the RMPC is both robustly recursively feasible and robustly stable. Finally, we discuss that the non-convex problem (3.18) to construct the hyper-rectangles $\mathscr{E}_k$ has a CP reformulation and an LP relaxation, and therefore can be solved in practice.

Algorithm 2 provides the event-based implementation policy of the RMPC method of this chapter.

Algorithm 2 Event-based implementation policy of RMPC

1: Calculate F according to (3.5), and K_i according to Definition 3.2.4, with $M < N-1$.
2: Initialize by setting $k = 0$, $j = 0$.
3: Measure x_0.
4: **loop**
5: Collect measurement x_{k+j} from sensors.
6: $k \leftarrow k+j$, $j \leftarrow 0$. ▷ Update
7: Solve $\mathscr{P}(x_k)$.
8: Solve Problem (3.18) for $j \in \mathsf{N}_{[N-1]}$ using Theorem 3.4.6 or 3.4.7.
9: Send the input sequence $\mathbf{U}^*_{k|k}$ to the actuators.
10: Send bounds $x^{*,p}_{k+j|k} - \underline{e}^p_{j,k}$ and $x^{*,p}_{k+j|k} + \overline{e}^p_{j,k}$ to the corresponding sensors, for all $j \in \mathsf{N}_{[N-1]}$.
11: Apply $u_k = u^*_{k|k}$ from the solution of $\mathscr{P}(x_k)$.
12: $j \leftarrow j+1$,
13: **if** $j > N-1$ **then**
14: **Go to** 6 ▷ Update triggered
15: **end if**
16: **for** $p \in \mathsf{N}_{[n]}$ **do**
17: Measure x^p_{k+j}.
18: **if** $x^p_{k+j} < x^{*,p}_{k+j|k} - \underline{e}^p_{j,k}$ or $x^p_{k+j} > x^{*,p}_{k+j|k} + \overline{e}^p_{j,k}$ **then**
19: **Go to** 6 ▷ Update triggered
20: **end if**
21: **end for**
22: Apply $u_k = u^*_{k+j|k}$ from the solution of $\mathscr{P}(x_k)$.
23: **Go to** 12 ▷ No update triggered
24: **end loop**

Remark 3.4.2 (Transmission protocol)**.** *We assume that all communication units on sensors, controller and actuators are clock-synchronized (by doing so, one can effectively reduce the listening time and conserve energy). At each instant k that the RMPC problem* (3.10) *is solved, the controller node sends: (i) $\mathbf{U}^*_{k|k}$ to the actuator nodes and (2) the state individual bounds $x^{*,p}_{k+j|k} - \underline{e}^p_{j,k}$ and $x^{*,p}_{k+j|k} + \overline{e}^p_{j,k}$ to the corresponding sensory nodes, for all $j \in \mathsf{N}_{[N-1]}$. We further assume that the n_x sensor units declare a triggering instant to each other, through a cost-efficient short-range transmission. Each individual sensor then declares its observed state to the controller.*

Theorem 3.4.3 (Event-based robust recursive feasibility)**.** *Consider the perturbed LTI dynamics* (3.2) *subject to the constraints and the disturbances given in* (3.3)*. Suppose that for some initial state $x_0 \in \mathbb{R}^{n_x}$, $\mathscr{P}(x_0)$ has a feasible solution. Then, the state and input trajectories of the dynamics* (3.2) *controlled by Algorithm 2 satisfy the constraints* (3.3)*, for all $k \in \mathbb{Z}_{\geq 0}$, i.e., robust recursive feasibility.*

Theorem 3.4.4 (Event-based robust convergence)**.** *Consider the perturbed LTI dynamics* (3.2) *subject to the constraints and the disturbances given in* (3.3)*. Suppose that for*

some initial state $x_0 \in \mathbb{R}^{n_x}$, $\mathscr{P}(x_0)$ has a feasible solution. Then, the state and input trajectories of dynamics (3.2)*, controlled by Algorithm 2, are such that $x_k \to \mathbb{T}_x$ and $u_k \to \mathbb{T}_u$, as $k \to \infty$, i.e., robust convergence.*

The proofs of Theorems 3.4.3 and 3.4.4 are provided in Subsections 3.5.3 and 3.5.4, respectively.

Remark 3.4.5 (Difference with standard analysis)**.** *In standard implementations of RMPC methods (including the RMPC method of this chapter), one is asked to guarantee that the application of $u^*_{k|k}$ on the plant does not render its following optimal control problem infeasible, recursively. Alternatively in an event-based implementation policy, one should ensure feasibility not only at the triggering instants but also between two consecutive triggering instants.*

The successful usage of the above results is conditioned upon the premise that there exist computationally tractable methods to construct the sets $\{\mathscr{E}_{j,k}\}_{j=1}^{N-1}$. We now revisit the optimization problem (3.18) to show that such a premise is valid by providing two frameworks: one in a CP form and another one in an LP form. Both of the frameworks are based on the same basis: the parametric-in-set constraints (3.18b)-(3.18f) can be reformulated into a new set of linear inequalities in terms of the vertices of each set $\mathscr{E}_{j,k}$. We shall call the polytope represented by the derived linear inequalities, the *principal* polytope $\bar{\mathscr{S}}$. Both frameworks try to find a maximum-volume hyper-rectangle $\mathscr{E}_{j,k}$ inscribed (or contained) in the principal polytope such that $0 \in \mathscr{E}_{j,k}$. This problem is closely related to a well-studied problem in the literature known as "inradius" of a polytopic set with respect to the polytopal norm induced by a hyper-rectangle with fixed (related to the LP form) or variable (related to the CP form) edge ratios. (See e.g., [76] and [77] for a detailed discussion on such problems.) In the LP framework, we partly employ some results from [77], see Subsection 3.5.5. We avoid reiterating the proofs of material borrowed from [77].

Theorem 3.4.6 (Volume maximization - CP reformulation)**.** *Consider a vector $\xi \in \mathbb{R}^p$, a matrix $M \in \mathbb{R}^{p\times k}$, and a polytope $\mathscr{S} = \{s \in \mathbb{R}^p : A_{\mathscr{S}} s \le b_{\mathscr{S}}\}$ containing the origin where $A_{\mathscr{S}} \in \mathbb{R}^{m\times p}$ and $b_{\mathscr{S}} \in \mathbb{R}^m$. The maximum volume hyper-rectangle $\mathscr{B}(l,u) \subset \mathbb{R}^k$ that contains the origin and satisfies $\xi \in \mathscr{S} \sim M\mathscr{B}(l,u)$ is $\mathscr{B}(-\underline{v}^*, \overline{v}^*)$ where $\underline{v}^*$ and $\overline{v}^*$ are the optimal solutions of the problem*

$$\begin{aligned} \min_{\underline{v},\overline{v}} \quad & f_q(\overline{v},\underline{v}) \\ \text{s.t.} \quad & \langle w^i, [\overline{v}^\top \ \underline{v}^\top]^\top \rangle \le b_{i,\mathscr{S}} - a_{i,\mathscr{S}}\xi, \text{ for all } i \in \mathsf{N}_{[m]}, \quad \overline{v} \ge 0, \ \underline{v} \ge 0, \end{aligned} \tag{3.19}$$

where for $q \in \{1,2\}$

$$f_1(\overline{v},\underline{v}) := -\sum_{j\in\mathsf{N}_{[k]}} \log(\overline{v}_j + \underline{v}_j), \tag{3.20a}$$

$$f_2(\overline{v},\underline{v}) := -\sum_{j\in\mathsf{N}_{[k]}} \log(\overline{v}_j) + \log(\underline{v}_j), \tag{3.20b}$$

and for all $j \in \mathsf{N}_{[k]}$

$$w_j^i = \begin{cases} \left(M^\top a_{i,\mathscr{S}}^\top\right)_j, & \text{if } \hat{w}_j^i = 1, \\ 0, & \text{otherwise}, \end{cases} \tag{3.21a}$$

$$w_{k+j}^i = \begin{cases} -\left(M^\top a_{i,\mathscr{S}}^\top\right)_j, & \text{if } \hat{w}_j^i = -1, \\ 0, & \text{otherwise}, \end{cases} \tag{3.21b}$$

with $\hat{w}^i := \operatorname{sign}(M^\top a_{i,\mathscr{S}}^\top)$, *for all* $i \in \mathsf{N}_{[m]}$.

Theorem 3.4.7 (Volume maximization - LP relaxation)**.** *Suppose the hypotheses in Theorem 3.4.6 hold.*

- (**q** = **1**) *The maximum volume r-constrained hyper-rectangle* $\mathscr{B}(l,u) \subset \mathbb{R}^k$ *that contains the origin and satisfies* $\xi \in \mathscr{S} \sim M\mathscr{B}(l,u)$ *is* $\mathscr{B}(z^*, z^* + \lambda^* r)$ *for which* $z^* \in \mathbb{R}^k$ *and* $\lambda^* \in \mathbb{R}$ *are the optimal solution of the problem*

$$\begin{aligned} \max_{z,\lambda} \quad & \lambda \\ \text{s.t.} \quad & A_{\mathscr{S}} M z + (A_{\mathscr{S}} M)^+ r\lambda \le b_{\mathscr{S}} - A_{\mathscr{S}}\xi \\ & z + \lambda r \ge 0,\ z \le 0, \end{aligned} \tag{3.22}$$

where for all $j \in \mathsf{N}_{[k]}$, *the* j*-th entry of* r *is defined as*

$$\begin{aligned} r_j := \quad \max_{z,\omega} \quad & \omega \\ \text{s.t.} \quad & A_{\mathscr{S}} M z \le b_{\mathscr{S}} - A_{\mathscr{S}}\xi \\ & A_{\mathscr{S}} M(z + \omega e_j) \le b_{\mathscr{S}} - A_{\mathscr{S}}\xi \\ & z + \omega e_j \ge 0,\ z \le 0, \end{aligned} \tag{3.23}$$

where $e_j \in \mathbb{R}^k$ *is the unit vector in the* j*-th direction and the polytope* $\bar{\mathscr{S}}$ *is*

$$\bar{\mathscr{S}} := \{z \in \mathbb{R}^k :\ A_{\mathscr{S}} M z \le b_{\mathscr{S}} - A_{\mathscr{S}}\xi\}.$$

- (**q** = **2**) *The maximum volume r-constrained hyper-rectangle* $\mathscr{B}(l,u) \subset \mathbb{R}^k$ *that contains the origin and satisfies* $\xi \in \mathscr{S} \sim M\mathscr{B}(l,u)$ *is* $\mathscr{B}(-\lambda^* r_1, \lambda^* r_2)$ *for which* $\lambda^* \in \mathbb{R}$ *is the optimal solution of the problem*

$$\begin{aligned} \max_{\lambda} \quad & \lambda \\ \text{s.t.} \quad & (W)^+ r\lambda \le B, \end{aligned} \tag{3.24}$$

where $r = \left(r_2^\top, r_1^\top\right)^\top$ *and for all* $j \in \mathsf{N}_{[2k]}$, *the* j*-th entry of* r *is defined as*

$$\begin{aligned} r_j := \quad \max_{\omega} \quad & \omega \\ \text{s.t.} \quad & W'(\omega e_j) \le B', \end{aligned} \tag{3.25}$$

where $e_j \in \mathbb{R}^{2k}$ is the unit vector in the j-th direction,

$$W = \left(w^1, \cdots, w^m\right)^\top, \qquad W' = \begin{pmatrix} W \\ -\mathsf{I}_k \quad 0_{k\times 1} \\ 0_{k\times 1} \quad -\mathsf{I}_k \end{pmatrix},$$

$$B = b_{\mathscr{S}} - A_{\mathscr{S}}\xi, \qquad B' = \left(B^\top, 0_{1\times 2k}\right)^\top,$$

and for all $i \in \mathsf{N}_{[m]}$, w^i are defined in (3.21).

Recall that the derived hyper-rectangle $\mathscr{B}$ using the CP or LP framework and the constraints $\xi \in \mathscr{S} \sim M\mathscr{B}(l,u)$ represent the hyper-rectangle $\mathscr{E}_{j,k}$ and the constraint (3.18b)-(3.18f), respectively. We should emphasize that although Theorem 3.4.6 or 3.4.7 provides a method to construct a set $\mathscr{E}_{j,k}$ with the maximal volume, the derived set is not unique (the cost functions related to both of the approaches are not strictly convex to guarantee the uniqueness of the optimal decision variables). In the remainder of this chapter, we denote the construction approach based on the CP (3.19) with $q = 1$ and $q = 2$ by CP_1 and CP_2, respectively. Furthermore, LP_1 represents the LP relaxation (3.22)-(3.23) of CP_1 and LP_2 denotes the LP relaxation (3.24)-(3.25) of CP_2.

Computational complexity: Consider the convex program (CP) $\{\min f_0(\eta),\ \text{s.t. } f_i(\eta) \le 0,\ \text{for all } i \in \mathsf{N}_{[n_c]}\}$ with the decision variable $\eta \in \mathbb{R}^{n_d}$, where the constraint functions $f_i : \mathbb{R}^{n_d} \to \mathbb{R}$ are convex. The computational effort to solve this convex program (or any optimization problem for that matter) depends on (i) the problem instance (i.e., the types of f_0 and the constraint functions f_i), (ii) the utilized algorithm (e.g., interior-point methods [75]), and (iii) the prescribed accuracy of the solution $\epsilon \in \mathbb{R}_{>0}$ (i.e., $f_0(x) - f^* \le \epsilon$ where f^* is the optimal value). Assuming these three items are set, one can then use the *arithmetic complexity* to provide a measure for the computational effort. In simple words, the arithmetic complexity provides an upper bound on the number of basic mathematical operations (that is called *flops*) required to attain an ε-close solution to the optimal value. (See e.g., [78] for an in-depth treatment of the complexity subject.) In what follows, we mainly concentrate on interior-point methods. These algorithms are a typical choice to solve moderate-size, smooth, convex problems by applying Newton's method [79, Chapter 9] to a sequence of equality constrained problems. (We strongly recommend the reader to consult with [79, Appendix C] and the references therein, which are specifically tailored for the complexity analysis of Newton's method.) Let us begin with providing the arithmetic complexity of a general LP $\{\max c^\top \eta :\ A\eta \le b\}$ where $A \in \mathbb{R}^{n_c \times n_d}$ and $b \in \mathbb{R}^{n_c}$. One can use for example the interior-point method proposed by Renegar [80] and achieve the total arithmetic complexity $\mathscr{O}\big((n_c + n_d)^{1.5} n_d^2 \ln(\frac{1}{\epsilon})\big)$. Suppose now $n_d \ll n_c$, the arithmetic complexity then becomes $\mathscr{O}\big(n_c^{1.5} n_d^2 \ln(\frac{1}{\epsilon})\big)$. One can even take a step further and use the mechanism introduced in [75, Subsection 3.4] and achieve the complexity $\mathscr{O}\big(n_c n_d^2 \ln(\frac{1}{\epsilon})\big)$. Now consider a general CP $\{\min f_0(x) :\ f_i(x) \le b_i, i \in \mathsf{N}_{[n_c]}\}$ where f_i are convex concordant functions [75, Chapter 2], for all $i \in \mathsf{Z}_{[n_c]}$. Assume further that (i) f_0 and the barrier function ϕ associated with the constraints are closed and self-concordant, (ii) the Hessian of $t f_0 + \phi$ is positive definite everywhere (see [79, Page 586]). One can use the barrier method introduced in [79, Subsection 11.3]. As a result, the arithmetic complexity of the algorithm is $\mathscr{O}\big(\sqrt{n_c}\ln(\frac{1}{\epsilon})\big)$.

It is important to remind the reader that this type analysis is in fact conservative. Roughly speaking, the actual behavior of the applied algorithm is highly affected by the structural properties of the constraints and the objective function. Nesterov and Nemirovski have used the notion of self-concordance functions to properly capture these properties [75, Chapter 2]. By doing so, they have been able to provide a unified framework for interior point methods.

Let us now explain what is the arithmetic complexity of the approaches proposed in Theorems 3.4.6 & 3.4.7. We adopt the following notion of an oracle to represent the convex problems in this paper. By doing so, one can provide the arithmetic complexity of the problems in Theorems 3.4.6 & 3.4.7 independent of the approach used. More precisely, we denote (i) by $\mathbf{lp}(n_c, n_d)$ the oracle complexity for solving $\max_{\eta} \{c^\top \eta : \ A\eta \leq b\}$ where $A \in \mathbb{R}^{n_c \times n_d}$, $c \in \mathbb{R}^{n_d}$ and $b \in \mathbb{R}^{n_c}$, and (ii) by $\mathbf{cp}(n_c, n_d)$ the oracle complexity for solving $\max_{\eta} \{f(\eta) : \ A\eta \leq b\}$ where the concave function $f : \mathbb{R}^{n_d} \to \mathbb{R}$, $A \in \mathbb{R}^{n_c \times n_d}$, and $b \in \mathbb{R}^{n_c}$. Based on the adopted complexity notion, we now provide the complexity of the approaches in Theorems 3.4.6 & 3.4.7.

Remark 3.4.8 (Oracle complexity)**.** ***CP reformulations:*** *The oracle complexity of problem* (3.19) *is* $\mathbf{cp}(m+2k, 2k)$. ***LP*$_1$ *relaxation (q=1):*** *The complexity of problem* (3.23) *is* $k \times \mathbf{lp}(2m+2k, k+1)$. *Moreover,* $\mathbf{lp}(m+2k, k+1)$ *represents the oracle complexity of problem* (3.22). ***LP*$_2$ *relaxation (q=2):*** *The complexity of problem* (3.25) *is* $2k \times \mathbf{lp}(m+2k, 1)$. *Furthermore,* $\mathbf{lp}(m, 1)$ *represents the oracle complexity of problem* (3.24).

A geometrical measure of sensitivity: As mentioned in Remark 3.4.1, we use Theorem 3.4.6 or 3.4.7 to solve the optimization problem (3.18), and find the hyper-rectangle $\mathscr{E}_{j,k}$. Suppose now $\mathscr{E}_{j,k}$ is found using either of the approaches proposed in Theorem 3.4.6 or 3.4.7. Two points of clarification are now in order. The first one has to do with the fact that a large volume of $\mathscr{E}_{j,k}$ and as a matter of fact even a large width of $\mathscr{E}_{j,k}$ along a coordinate axis p (that is $\underline{e}^p_{j,k} + \overline{e}^p_{j,k}$) can be misleading. In doing so, let us define two types of 2-norm balls. Suppose that the polytope that represents the constraint (3.18b)-(3.18f) is denoted by $\overline{\mathscr{S}} \subset \mathbb{R}^{n_x}$. We denote by $\mathscr{B}^2_{\mathrm{c},\overline{\mathscr{S}}}(x_\mathrm{c}, r_\mathrm{c})$ the maximal 2-norm ball that fits inside $\overline{\mathscr{S}}$. The center x_c is the so-called *Chebychev center* of the polytope $\overline{\mathscr{S}}$. (See e.g., [79, Chapters 4&8] for more details on the Chebychev center of a polyhedron.) Hence, $\mathscr{B}^2_{\mathrm{c},\overline{\mathscr{S}}} := \{x_\mathrm{c} + y \in \mathbb{R}^{n_x} : \ \|y\| \leq r_\mathrm{c}\} \subseteq \overline{\mathscr{S}}$. We further denote by $\mathscr{B}^2_{\mathrm{o},\overline{\mathscr{S}}}(r_\mathrm{o})$ the maximal 2-norm ball that is centered at the origin and $\mathscr{B}^2_{\mathrm{o},\overline{\mathscr{S}}}(r_\mathrm{o}) \subseteq \overline{\mathscr{S}}$. Thus, $\mathscr{B}^2_{\mathrm{o},\overline{\mathscr{S}}}(r_\mathrm{o}) := \{y \in \mathbb{R}^{n_x} : \ \|y\| \leq r_\mathrm{o}\} \subseteq \overline{\mathscr{S}}$. We now introduce the measure $\frac{r_\mathrm{c}}{r_\mathrm{o}}$ of $\overline{\mathscr{S}}$ that characterizes an important feature of the derived hyper rectangle $\mathscr{E}_{j,k}$ using the frameworks proposed in Theorems 3.4.6 & 3.4.7. It is worth mentioning that (i) this property of $\overline{\mathscr{S}}$ is inherited from the constraints (3.18b)-(3.18f), and (ii) the ratio $\frac{r_\mathrm{c}}{r_\mathrm{o}} \geq 1$ (the equality case happens if $\overline{\mathscr{S}}$ is symmetric with respect to the origin). The second point of clarification is related to the fact that the hyper-rectangles constructed based on the approaches CP_2 and LP_2 are more symmetric with respect to the origin compared to the ones constructed based on the approaches CP_1 and LP_1.

Remark 3.4.9 (Directional sensitivity to prediction errors)**.** ***Well-shaped case:*** *If the ratio* $\frac{r_c}{r_\mathrm{o}} \sim 1$*, then, the optimal hyper-rectangle* $\mathscr{E}_{j,k}$ *derived based on Theorem 3.4.6 or 3.4.7*

is relatively symmetric with respect to the origin. As a result, the difference between the lower-bound $\underline{e}^p_{j,k}$ and the upper-bound $\overline{e}^p_{j,k}$ is not significant (i.e., the triggering mechanism shows a similar behavior irrespective of the direction of error mismatch that occurs along the p-th coordinate). See Figure 3.1a. ***Ill-shaped case:*** *If the ratio $\frac{r_c}{r_\circ} \gg 1$, then, the optimal hyper-rectangle $\mathscr{E}_{j,k}$ derived based on Theorem 3.4.6 or 3.4.7 is relatively asymmetric with respect to the origin. As a result, the difference between the lower-bound $\underline{e}^p_{j,k}$ and the upper-bound $\overline{e}^p_{j,k}$ can become significant (i.e., the triggering mechanism shows an extreme level of sensitivity to the direction of the prediction error along the coordinate $p \in \mathsf{N}_{[n_x]}$ with a large $|\overline{e}^p_{j,k} - \underline{e}^p_{j,k}|$). See Figure 3.1b.*

3.5. TECHNICAL PROOFS

3.5.1. PROOF OF THEOREM 3.3.4

(Proof by induction) It is postulated by the theorem that there is a feasible solution for $\mathscr{P}(x_0)$. For any $k \geq 1$, we then have to show that given a feasible solution for $\mathscr{P}(x_k)$ and applying the first element of the sequence $\mathbf{U}^*_{k|k}$, there exits a feasible candidate solution for the subsequent problem $\mathscr{P}(x_{k+1})$, for all $w_k \in \mathscr{W}$. The proof of claim (ii) immediately follows from this feasibility result. Because feasible solutions for $\mathscr{P}(x_k)$ imply the satisfaction of the constraints on x_k and u_k in (3.3), through (3.10d) and (3.10e) with $i = 0$.

Suppose k is the instant at which the problem $\mathscr{P}(x_k)$ has been solved, with its corresponding optimal input and state trajectories $\mathbf{U}^*_{k|k}$ and $\mathbf{X}^*_{k|k}$, respectively. Evidently, the constraints (3.10c)-(3.10f) are satisfied. Define $A_{\mathrm{cl}} := (A + BF)$. At the instant $k+1$, it is trivial to show that the disturbance can be derived from the most recent measurement, i.e., $w_k = x_{k+1} - x^*_{k+1|k}$. Then, a candidate control sequence $\hat{\mathbf{U}}_{k+1|k+1}$ is

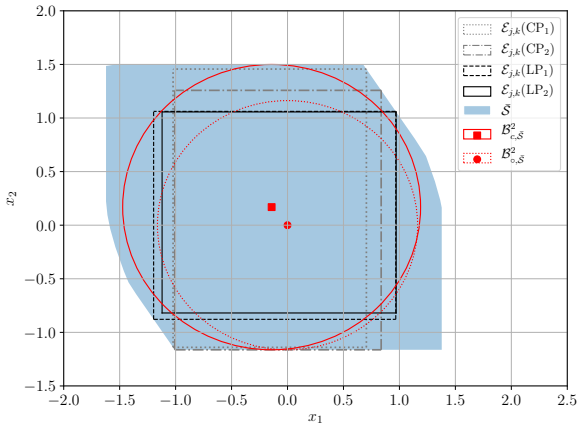
$$\begin{pmatrix} \hat{u}_{k+1|k+1} \\ \hat{u}_{k+2|k+1} \\ \vdots \\ \hat{u}_{k+N-1|k+1} \\ \hat{u}_{k+N|k+1} \end{pmatrix} = \begin{pmatrix} u^*_{k+1|k} \\ u^*_{k+2|k} \\ \vdots \\ u^*_{k+N-1|k} \\ Fx^*_{k+N|k} \end{pmatrix} + \begin{pmatrix} K_0 L_0 \\ K_1 L_1 \\ \vdots \\ K_{N-2} L_{N-2} \\ F L_{N-1} \end{pmatrix} w_k, \tag{3.26a}$$

which results into the candidate state trajectory $\hat{\mathbf{X}}_{k+1|k+1}$

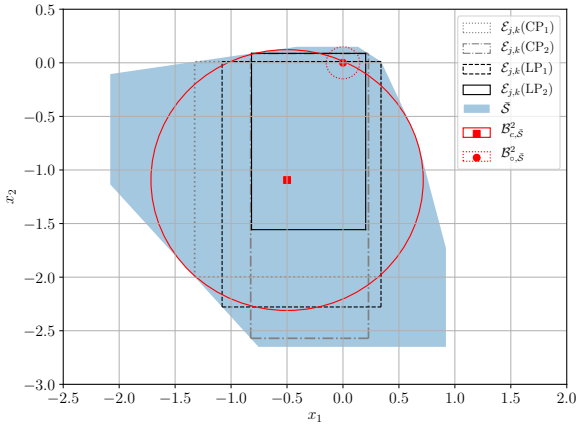
$$\begin{pmatrix} \hat{x}_{k+1|k+1} \\ \hat{x}_{k+2|k+1} \\ \vdots \\ \hat{x}_{k+N|k+1} \\ \hat{x}_{k+N+1|k+1} \end{pmatrix} = \begin{pmatrix} x^*_{k+1|k} \\ x^*_{k+2|k} \\ \vdots \\ x^*_{k+N|k} \\ A_{\mathrm{cl}} \hat{x}_{k+N|k+1} \end{pmatrix} + \begin{pmatrix} L_0 \\ L_1 \\ \vdots \\ L_{N-1} \\ 0 \end{pmatrix} w_k. \tag{3.26b}$$

We now establish that the candidate trajectories in (3.26) satisfy the constraints (3.10b)-(3.10f). This in turn implies that $\tilde{\mathbf{U}}_{k+1|k+1}$ is a feasible solution for $\mathscr{P}(x_{k+1})$, if x_{k+1} is given by the system dynamics (3.2).

It simply holds that $\hat{x}_{k+1|k+1} = x^*_{k+1|k} + L_0 w_k = x^*_{k+1|k} + w_k = x_{k+1|k+1}$. Hence, the *initial state* constraint (3.10b) is satisfied. Making use of the linearity of dynamics and



(a) Well-shaped polytope example, $\frac{r_c}{r_o} = 1.0861$.



(b) Ill-shaped polytope example, $\frac{r_c}{r_o} = 8.0669$.

Figure 3.1: Comparison of the CP and LP approaches in Theorems 3.4.6 & 3.4.7 to compute an inner hyper-rectangle $\mathcal{E}_{j,k}$ inside a given polytope $\overline{\mathcal{S}}$. **(Well-shaped example)** In Figure 3.1a, the Chebychev center x_c is relatively close to the origin (the area of $\overline{\mathcal{S}}$ is distributed in a somewhat uniform manner around the origin). All the four approaches provide close behaviors. As a result, the directional sensitivity of $\mathcal{E}_{j,k}$ for all approaches is low. **(Ill-shaped example)** Contrary to the previous case, the Chebychev center x_c of $\overline{\mathcal{S}}$ is relatively far from the origin in Figure 3.1b. The hyper-rectangles $\mathcal{E}_{j,k}$ derived based on the approaches CP_2 and LP_2 provide a more symmetric result compared to the ones computed based on the approaches CP_1 and LP_1.

the definition (3.6e), we have for the states further in the sequence

$$\hat{x}_{k+2|k+1} = A\underbrace{(x^*_{k+1|k} + L_0 w_k)}_{=\hat{x}_{k+1|k+1}} + B\underbrace{(u^*_{k+1|k} + K_0 L_0 w_k)}_{=\hat{u}_{k+1|k+1}} = x^*_{k+2|k} + \underbrace{L_1}_{=(A+BK_0)L_0} w_k,$$

$$\hat{x}_{k+3|k+1} = A\underbrace{(x^*_{k+2|k} + L_1 w_k)}_{=\hat{x}_{k+2|k+1}} + B\underbrace{(u^*_{k+2|k} + K_1 L_1 w_k)}_{=\hat{u}_{k+2|k+1}} = x^*_{k+3|k} + \underbrace{L_2}_{=(A+BK_1)L_1} w_k,$$

....

We lastly need to show that the terminal state of $\hat{\mathbf{X}}_{k+1|k+1}$ satisfies the nominal system dynamics. To do so,

$$\hat{x}_{k+N+1|k+1} = A\underbrace{(x^*_{k+N|k} + L_{N-1} w_k)}_{=\hat{x}_{k+N|k+1}} + BF\underbrace{(x^*_{k+N|k} + L_{N-1} w_k)}_{=\hat{x}_{k+N|k+1}} = A_{cl}\hat{x}_{k+N|k+1}.$$

As a result, the *dynamics* constraints (3.10c) are satisfied by $\hat{\mathbf{U}}_{k+1|k+1}$ and $\hat{\mathbf{X}}_{k+1|k+1}$.

Since $\mathbf{U}^*_{k|k}$ is a feasible solution for $\mathscr{P}(x_k)$, $u^*_{k+i+1|k} \in \mathscr{U}_{i+1}$, $\forall i \in \mathsf{Z}_{[N-2]}$. We also have $\mathscr{U}_{i+1} = \mathscr{U}_i \sim K_i L_i \mathscr{W}$ by the relation (3.6a). Using Definition 3.2.2, i.e., the Pontryagin difference, it thus follows that $\hat{u}_{k+i+1|k+1} = u^*_{k+i+1|k} + K_i L_i w_k \in \mathscr{U}_i$, $\forall i \in \mathsf{Z}_{[N-2]}$. We next need to show that $\hat{u}_{k+N|k+1} \in \mathscr{U}_{N-1}$. Considering $x^*_{k+N|k} \in \mathscr{X}_f$, the relation (3.8) implies that $\hat{x}_{k+N|k+1} = x^*_{k+N|k} + L_{N-1} w_k \in \mathscr{R}$. We then have from the relation (3.7d) that $\hat{u}_{k+N|k+1} = F\hat{x}_{k+N|k+1} \in \mathscr{U}_{N-1}$. As a result, $\hat{u}_{k+i+1|k+1} \in \mathscr{U}_i$, for all $i \in \mathsf{Z}_{[N-1]}$ and thereby the candidate input trajectory satisfies the constraints (3.10d).

With a similar argument, one can show that the state predictions resulting from the candidate trajectory are also inside their respective sets. From $x^*_{k+i+1|k} \in \mathscr{X}_{i+1}$, $\forall i \in \mathsf{Z}_{[N-2]}$ and the definition of the Pontryagin difference, $\hat{x}_{k+i+1|k+1} = x^*_{k+i+1|k} + L_i w_k \in \mathscr{X}_i, \forall i \in \mathsf{Z}_{[N-2]}$. Since $x^*_{k+N|k} \in \mathscr{X}_f$, the relation (3.8) gives $\hat{x}_{k+N|k+1} \in \mathscr{R}$. Then, the relation (3.7b) implies that $\hat{x}_{k+N|k+1} \in \mathscr{X}_{N-1}$. Hence, $\hat{x}_{k+i+1|k+1} \in \mathscr{X}_i$, $\forall i \in \mathsf{Z}_{[N-1]}$. The constraints (3.10e) are satisfied by the candidate solution. Lastly, we are required to show that the terminal state $\hat{x}_{k+N+1|k+1} \in \mathscr{X}_f$, i.e., the satisfaction of the relation (3.10f). This simply follows from the previous observations and the relation (3.7a),

$$\hat{x}_{k+N|k+1} \in \mathscr{R} \underset{(3.7a)}{\Rightarrow} A_{cl}\hat{x}_{k+N|k+1} + L_{N-1} w_k \in \mathscr{R}, \forall w_k \in \mathscr{W} \underset{(3.8)}{\Rightarrow} \hat{x}_{k+N+1|k+1} \in \mathscr{X}_f.$$

The proof of claim (i) and as a result the proof of claim (ii) are concluded.

3.5.2. PROOF OF THEOREM 3.3.5

Recall the candidate trajectories $\hat{\mathbf{U}}_{k+1|k+1}$ and $\hat{\mathbf{X}}_{k+1|k+1}$ as defined in (3.26a) and (3.26b), respectively, for $\mathscr{P}(x_{k+1})$. In a nutshell, this proof consists of deriving a positive upper bound for $J(x_k, \mathbf{U}^*_{k|k}) - J(x_{k+1}, \hat{\mathbf{U}}_{k+1|k+1})$ that decreases as $k \to \infty$. Notice that the optimal solution for $\mathscr{P}(x_{k+1})$ has a cost that is smaller than or equal to the cost of the candidate solution , i.e., $J(x_{k+1}, \mathbf{U}^*_{k+1|k+1}) \leq J(x_{k+1}, \hat{\mathbf{U}}_{k+1|k+1})$.

By virtue of Lemma 3.2.3, one can observe that, for all $i \in \mathsf{Z}_{[N-2]}$ and for all $w_k \in \mathscr{W}$,

$$\begin{aligned} d(\hat{u}_{k+i+1|k+1}, \mathscr{T}_{u,i}, R) &\leq d(u^*_{k+i+1|k}, \mathscr{T}_{u,i} \sim K_i L_i \mathscr{W}, R) \\ &= d(u^*_{k+i+1|k}, \mathscr{T}_{u,i+1}, R), \\ d(\hat{x}_{k+i+1|k+1}, \mathscr{T}_{x,i}, Q) &\leq d(x^*_{k+i+1|k}, \mathscr{T}_{x,i} \sim L_i \mathscr{W}, Q) \\ &= d(x^*_{k+i+1|k}, \mathscr{T}_{x,i+1}, Q). \end{aligned}$$

Notice that $x^*_{k+N|k} \in \mathscr{X}_f$ from (3.10f). The relation (3.7c) implies that $\hat{x}_{k+N|k+1} \in \mathscr{T}_{x,N-1}$. Moreover, from (3.7e) it follows that $\hat{u}_{k+N|k+1} \in \mathscr{T}_{u,N-1}$. Thus, the costs associated to $\hat{x}_{k+N|k+1}$ and $\hat{u}_{k+N|k+1}$ become zero for the candidate sequences. This last observation enables us to determine an upper bound for the cost related to the candidate solution as follows,

$$\begin{aligned} &J(x_{k+1}, \hat{\mathbf{U}}_{k+1|k+1}) = \\ &\sum_{i=0}^{N-1} d(\hat{x}_{k+i+1|k+1}, \mathscr{T}_{x,i}, Q) + d(\hat{u}_{k+i+1|k+1}, \mathscr{T}_{u,i}, R) \\ &\quad \leq \sum_{i=0}^{N-2} d(x^*_{k+i+1|k}, \mathscr{T}_{x,i+1}, Q) + d(u_{k+i+1|k}, \mathscr{T}_{u,i+1}, R) \\ &\quad = J(x_k, \mathbf{U}^*_{k|k}) - d(x^*_{k|k}, \mathscr{T}_{x,0}, Q) - d(u^*_{k|k}, \mathscr{T}_{u,0}, R), \end{aligned}$$

as well as an upper bound on the optimal cost for the optimization performed at $k+1$, that is

$$\begin{aligned} J(x_{k+1}, \mathbf{U}^*_{k+1|k+1}) &\leq J(x_{k+1}, \hat{\mathbf{U}}_{k+1|k+1}) \\ &\leq J(x_k, \mathbf{U}^*_{k|k}) - d(x^*_{k|k}, \mathscr{T}_{x,0}, Q) - d(u^*_{k|k}, \mathscr{T}_{u,0}, R). \end{aligned}$$

Since the distance function d is nonnegative, it holds that $J(x_k, \mathbf{U}_{k|k}) \geq 0$. It follows from the above relation that $J(x_k, \mathbf{U}^*_k)$ decreases with increasing k and converges to a steady value. This fact in turn implies $d(x^*_{k|k}, \mathscr{T}_{x,0}, Q) + d(u^*_{k|k}, \mathscr{T}_{u,0}, R) \to 0$, as $k \to \infty$, from which we can conclude that $x_k \to \mathbb{T}_x$ and $u_k \to \mathbb{T}_u$ as $k \to \infty$. The claim is proved.

3.5.3. PROOF OF THEOREM 3.4.3

Consider k denotes the last instant at which the optimization problem $\mathscr{P}(x_k)$ is solved. Let $j \in \mathsf{N}_{[N-2]}$. Suppose without loss of generality that the mechanism is enabled at the instant $k+j+1$, i.e., $e_{k+j+1|k} \notin \mathscr{E}_{j+1,k}$. Hence, for all $j \in \mathsf{N}_{[j]}$, $e_{k+j|k} \in \mathscr{E}_{j,k}$.

(*Inter-event feasibility*) We first define the *candidate* input and state trajectories

$$\begin{aligned} \tilde{\mathbf{U}}_{k+j|k+j} &:= \{\tilde{u}_{k+j+i|k+j}\}_{i=0}^{N-1}, \\ \tilde{\mathbf{X}}_{k+j|k+j} &:= \{\tilde{x}_{k+j+i|k+j}\}_{i=0}^{N}, \end{aligned}$$

as follows,

$$\tilde{u}_{k+j+i|k+j} := \begin{cases} u^*_{k+j+i|k} + \tilde{K}_i \tilde{L}_i e_{k+j|k}, & \text{for } i \in \mathsf{Z}_{[N-j-1]}, \\ F(A+BF)^{(j+i-N)} x^*_{k+N|k} + \tilde{K}_i \tilde{L}_i e_{k+j|k}, & \text{for } i \in \mathsf{Z}_{[N-1]} / \mathsf{Z}_{[N-j-1]}, \end{cases} \tag{3.27a}$$

$$\tilde{x}_{k+j+i|k+j} := \begin{cases} x^*_{k+j+i|k} + \tilde{L}_i e_{k+j|k}, & \text{for } i \in \mathsf{Z}_{[N-j]}, \\ (A+BF)^{(j+i-N)} x^*_{k+N|k} + \tilde{L}_i e_{k+j|k}, & \text{for } i \in \mathsf{Z}_{[N]}/\mathsf{Z}_{[N-j]}, \end{cases} \tag{3.27b}$$

where the prediction error $e_{k+j|k}$, the disturbance gains $\tilde{K}_i$, and the transition matrices $\tilde{L}_i$ are given in (3.12), (3.16a), and (3.16b), respectively. We now establish that these candidate trajectories satisfy the constrains (3.10b)-(3.10f). By doing so, $x_{k+j} \in \mathbb{X}$ for all j such that $e_{k+j|k} \in \mathscr{E}_{j,k}$, i.e., the state x_{k+j} does not violate the constraint (3.3). Observe that $\tilde{x}_{k+j|k+j} = x_{k+j}$ and $\tilde{u}_{k+j|k+j} = u^*_{k+j|k}$ are the actual, applied input and the actual, observed state at the instant $k+j$, respectively. It is not difficult to see that the equality constraints (3.10b)-(3.10c) are satisfied given the initial state $\tilde{x}_{k+j|k+j}$ and the candidate input trajectory $\tilde{U}_{k+j|k+j}$ (3.27a) (simply by successive application of the nominal dynamics(3.4)). We next deduce that the constraints (3.10d)-(3.10e) are satisfied. Since $e_{k+j|k} \in \mathscr{E}_{j,k}$, the constraints (3.18c)-(3.18d) imply that $x^*_{k+j+i|k} \in \mathscr{X}_i \sim \tilde{L}_i \mathscr{E}_{j,k}$ and $u^*_{k+j+i|k} \in \mathscr{U}_i \sim \tilde{K}_i \tilde{L}_i \mathscr{E}_{j,k}$, for all $i \in \mathsf{Z}_{[N-1]}$. By virtue of (3.27) for all $i \in \mathsf{Z}_{[N-1]}$ and the definition of the Pontryagin difference (see Definition 3.2.2), $\tilde{x}_{k+j+i|k+j} = x^*_{k+j+i|k} + \tilde{L}_i e_{k+j|k} \in \mathscr{X}_i$ and $\tilde{u}_{k+j+i|k+j} = u^*_{k+j+i|k} + \tilde{K}_i \tilde{L}_i e_{k+j|k} \in \mathscr{U}_i$ (where the convention (3.13) is used to simplify the notation). At last, we show that the constraint (3.10f) holds. The disturbance gains K_i are designed to render the nominal dynamics (3.4) nilpotent in M steps where $M < N-1$ steps. This nilpotency in turn implies that $L_{N-1} = 0_{n\times n}$. Thus, $\mathscr{X}_f = \mathscr{R}$ from the terminal set definition (3.8) and the final transition matrix $\tilde{L}_N = L_{N-1} \cdot A = 0_{n_x \times n_x}$ is implied by the construction (3.16b). Considering the relation (3.7a) and $L_{N-1} = 0_{n_x \times n_x}$, we arrive at $(A+BF)^j x^*_{k+N|k} \in \mathscr{R} = \mathscr{X}_f$. As a result, $\tilde{x}_{k+j+N|k+j} = (A+BF)^j x^*_{k+N|k} + \tilde{L}_N e_{k+j|k} \in \mathscr{X}_f$, i.e., the constraint (3.10f) is respected by the candidate solution $\tilde{x}_{k+j+N|k+j}$. Based on the arguments provided above, we established that (i) the state $x_{k+j} \in \mathbb{X}$ if $e_{k+j|k} \in \mathscr{E}_{j,k}$, and (ii) there exist the candidate trajectories $\tilde{\mathbf{U}}_{k+j|k+j}$ and $\tilde{\mathbf{X}}_{k+j|k+j}$ that satisfy all the constraints (3.10b)-(3.10f) of the RMPC. In other words, we showed that the input sequence $\{u^*_{k+i|k}\}_{i=1}^{j}$ and the state sequences $\{x_{k+i}\}_{i=1}^{j}$ both remain feasible if $e_{k+i|k} \in \mathscr{E}_{i,k}$, for all $i \in \mathsf{N}_j$.

(*Event recursive feasibility*) In the next step, we should guarantee the optimization problem $\mathscr{P}(x_{k+j+1})$ is feasible in the case a triggering is enabled at the instant $k+j+1$. To this end, we employ the introduced candidate trajectories to show the Problem (3.10) is feasible at the triggering instant $k+j+1$. We now define two *new* candidate trajectories $\hat{U}_{k+j+1|k+j+1} := \{\hat{u}_{k+j+1+i|k+j+1}\}_{i=0}^{N-1}$ and $\hat{X}_{k+j+1|k+j+1} := \{\hat{x}_{k+j+1+i|k+j+1}\}_{i=0}^{N}$,

$$\hat{u}_{k+j+1+i|k+j+1} := \begin{cases} \tilde{u}_{k+j+1+i|k+j} + K_i L_i w_{k+j}, & \text{for } i \in \mathsf{Z}_{[N-2]}, \\ F\tilde{x}_{k+j+N|k+j} + FL_{N-1} w_{k+j}, & \text{for } i = N-1, \end{cases} \tag{3.28a}$$

$$\hat{x}_{k+j+1+i|k+j+1} := \begin{cases} \tilde{x}_{k+j+1+i|k+j} + L_i w_{k+j}, & \text{for } i \in \mathsf{Z}_{[N-1]}, \\ (A+BF)\hat{x}_{k+j+N|k+j}, & \text{for } i = N, \end{cases} \tag{3.28b}$$

where $w_{k+j} \in \mathscr{W}$ is the perturbation at the instant $k+j$.

We now show that the candidate trajectories $\hat{U}_{k+j+1|k+j+1}$ and $\hat{X}_{k+j+1|k+j+1}$ satisfy the constraints (3.10b)-(3.10f), i.e., the feasibility of $\mathscr{P}(x_{k+j+1})$. Recall that $\tilde{x}_{k+j|k+j} =$

x_{k+j} and $\tilde{u}_{k+j|k+j} = u^*_{k+j|k}$. Then, $x_{k+j+1} = Ax_{k+j} + Bu^*_{k+j|k} + w_{k+j}$ can be rewritten as $x_{k+j+1} = A\tilde{x}_{k+j|k+j} + B\tilde{u}_{k+j|k+j} + w_{k+j}$. Observe that $\tilde{x}_{k+j+1|k+j} = A\tilde{x}_{k+j|k+j} + B\tilde{u}_{k+j|k+j}$ and also that $L_0 = \mathsf{I}_{n_x}$ from the definition (3.6e). Since $x_{k+j+1} = \tilde{x}_{k+j+1|k+j} + L_0 w_{k+j}$, the constraint (3.10b) holds. Moreover, it follows that the remaining entries in $\hat{X}_{k+j+1|k+j+1}$ respect the equality constraints(3.10c) given the initial state $\hat{x}_{k+j+1|k+j+1}$ and the input trajectory $\hat{U}_{k+j+1|k+j+1}$. We next study the satisfaction of the constraints (3.10d)-(3.10e). Recall that $\tilde{x}_{k+j+1+i|k+j} \in \mathscr{X}_{i+1}$, for all $i \in \mathsf{Z}_{[N-2]}$. Considering the constraint tightening (3.6b) ($\mathscr{X}_{i+1} = \mathscr{X}_i \sim L_i\mathscr{W}$), it is not difficult to see that $\hat{x}_{k+j+1+i|k+j+1} = \tilde{x}_{k+j+1+i|k+j} + L_i w_{k+j} \in \mathscr{X}_i$, for all $i \in \mathsf{Z}_{[N-2]}$. For $i = N-1$, we have $\tilde{x}_{k+j+N|k+j} \in \mathscr{X}_f$. Moreover, $\mathscr{X}_f = \mathscr{R}$ since $L_{N-1} = 0_{n_x \times n_x}$ (from the nilpotency). Hence, $\tilde{x}_{k+j+N|k+j} \in \mathscr{X}_{N-1}$, and as a result, $\hat{x}_{k+j+N|k+j+1} = \tilde{x}_{k+j+N|k+j} + L_{N-1} w_{k+j} \in \mathscr{X}_{N-1}$. For all $i \in \mathsf{Z}_{[N-2]}$, $\tilde{u}_{k+j+1+i|k+j} \in \mathscr{U}_{i+1}$ where $\mathscr{U}_{i+1} = \mathscr{U}_i \sim K_i L_i \mathscr{W}$ from the constraint tightening (3.6a). Thus, $\hat{u}_{k+j+1+i|k+j+1} = \tilde{u}_{k+j+1+i|k+j} + K_i L_i w_{k+j} \in \mathscr{U}_i$, for all $i \in \mathsf{Z}_{[N-2]}$. It remains to show that $\hat{u}_{k+j+N|k+j+1} \in \mathscr{U}_{N-1}$. Notice that $\tilde{x}_{k+j+N|k+j} \in \mathscr{X}_f$. Considering $L_{N-1} = 0_{n_x \times n_x}$ (i.e., the nilpotency condition) and the relation (3.7d), it follows that $\hat{u}_{k+j+N|k+j+1} = F\tilde{x}_{k+j+N|k+j} \in \mathscr{U}_{N-1}$. As a result, the conditions (3.10d)-(3.10e) hold. Lastly, it needs to be shown that the terminal state $\hat{x}_{k+j+1+N|k+j+1} \in \mathscr{X}_f$, i.e., the terminal constraint (3.10f). We have $\hat{x}_{k+j+N|k+j+1} \in \mathscr{X}_{N-1}$. Thus, $\hat{x}_{k+j+1+N|k+j+1} = (A+BF)\hat{x}_{k+j+N|k+j+1} \in \mathscr{R} = \mathscr{X}_f$ (since $L_{N-1} = 0_{n_x \times n_x}$). The claim of the theorem holds.

3.5.4. PROOF OF THEOREM 3.4.4

In the first phase (*inter-triggering cost function decay*), we show that one can utilize the trajectories $\tilde{\mathbf{U}}_{k+j|k+j}$ and $\tilde{\mathbf{X}}_{k+j|k+j}$, given in (3.27), to define the cost $J(x_{k+j}, \tilde{\mathbf{U}}_{k+j|k+j})$ such that $J(x_{k+j}, \tilde{\mathbf{U}}_{k+j|k+j}) \leq J(x_k, \mathbf{U}^*_{k|k})$. In the second phase (*triggering cost function decay*), we demonstrate that one instead use the secondary trajectories $\hat{\mathbf{U}}_{k+j+1|k+j+1}$ and $\hat{\mathbf{X}}_{k+j+1|k+j+1}$, given in (3.28), to define the cost

$$J(x_{k+j+1}, \hat{\mathbf{U}}_{k+j+1|k+j+1}),$$

such that $J(x_{k+j+1}, \hat{\mathbf{U}}_{k+j+1|k+j+1}) \leq J(x_k, \mathbf{U}^*_{k|k})$.

(*Inter-triggering cost function decay*) Recall that for all $i \in \mathsf{Z}_{[N]}$ and for all $j \in \mathsf{N}_{[N-1]}$: (i) using the convention (3.14) for $j+i \in \mathsf{N}_{[N-1]}$, $s^*_{x,k+j+i|k}$ and $s^*_{u,k+j+i|k}$ are the points inside the target sets $\mathscr{T}_{x,j+i}$ and $\mathscr{T}_{u,j+i}$ that have the minimal distance to $x^*_{k+j+i|k}$ and $u^*_{k+j+i|k}$, respectively; and (ii) using the convention (3.15) for $j+i \in \mathsf{N}_{[N-1+j]}/\mathsf{N}_{[N-1]}$, $s^*_{x,k+j+i|k} = x^*_{k+j+i|k}$ and $s^*_{u,k+j+i|k} = u^*_{k+j+i|k}$. Let $k+j$ be the instant at which the triggering mechanism is being evaluated where $j \in \mathsf{N}_{[N-1]}$. In a nutshell, we show that if $e_{k+j|k} \in \mathscr{E}_{j,k}$, then, $J(x_{k+j}, \tilde{\mathbf{U}}_{k+j|k+j}) - J(x_k, \mathbf{U}^*_{k|k}) \leq 0$.

The definition (3.27b) implies that

$$d(\tilde{x}_{k+j+i|k+j}, \mathscr{T}_{x,i}, Q) = d(x^*_{k+j+i|k} + \tilde{L}_i e_{k+j|k}, \mathscr{T}_{x,i}, Q).$$

Observe that $e_{k+j|k} \in \mathscr{E}_{j,k}$ is provided by the triggering mechanism. A direct application of Lemma 3.2.3 leads to $d(\tilde{x}_{k+j+i|k+j}, \mathscr{T}_{x,i}, Q) \leq d(x^*_{k+j+i|k}, \mathscr{T}_{x,i} \sim \tilde{L}_i \mathscr{E}_{j,k}, Q)$. Furthermore, the relation (3.18e) guarantees that the point $s^*_{x,k+j+i|k} \in \mathscr{T}_{x,i} \sim \tilde{L}_i \mathscr{E}_{j,k}$, hence,

$d(x^*_{k+j+i|k}, \mathcal{T}_{x,i} \sim \tilde{L}_i \mathscr{E}_{j,k}, Q) \leq d(x^*_{k+j+i|k}, s^*_{x,k+j+i|k}, Q)$. All in all,

$$d(\tilde{x}_{k+j+i|k+j}, \mathcal{T}_{x,i}, Q) \leq d(x^*_{k+j+i|k}, s^*_{x,k+j+i|k}, Q). \tag{3.29}$$

a) Suppose now $j+i \in \mathsf{N}_{[N-1]}$. From the definition (3.14), $d(x^*_{k+j+i|k}, s^*_{x,k+j+i|k}, Q) = d(x^*_{k+j+i|k}, \mathcal{T}_{x,j+i}, Q)$. Thus, we conclude that

$$d(\tilde{x}_{k+j+i|k+j}, \mathcal{T}_{x,i}, Q) \leq d(x^*_{k+j+i|k}, \mathcal{T}_{x,j+i}, Q).$$

One can follow a similar path and arrive at, for all $j+i \in \mathsf{Z}_{[N-1]}$, $d(\tilde{u}_{k+j+i|k+j}, \mathcal{T}_{u,i}, R) \leq d(u^*_{k+j+i|k}, \mathcal{T}_{u,j+i}, R)$. Up until this point, we demonstrated that

$$\sum_{i=0}^{N-1-j} d(\tilde{x}_{k+j+i|k+j}, \mathcal{T}_{x,i}, Q) + d(\tilde{u}_{k+j+i|k+j}, \mathcal{T}_{u,i}, R) \leq \sum_{i=0}^{N-1-j} d(x^*_{k+j+i|k+j}, \mathcal{T}_{x,j+i}, Q) + d(u^*_{k+j+i|k}, \mathcal{T}_{u,j+i}, R). \tag{3.30}$$

b) Suppose next $j+i \in \mathsf{N}_{[N-1+j]} / \mathsf{N}_{[N-1]}$. Notice that we use the convention (3.15) for $s^*_{x,k+j+i|k}$ and $s^*_{u,k+j+i|k}$, for $j+i \geq N$. Based on the relation (3.29), it is not difficult to infer that $d(\tilde{x}_{k+j+i|k+j}, \mathcal{T}_{x,i}, Q) = 0$. In a similar fashion, one can deduce that $d(\tilde{u}_{k+j+i|k}, \mathcal{T}_{u,i}, R) = 0$. We then have

$$\sum_{i=N-j}^{N-1} d(\tilde{x}_{k+j+i|k}, \mathcal{T}_{x,i}, Q) + d(\tilde{u}_{k+j+i|k}, \mathcal{T}_{u,i}, R) = 0. \tag{3.31}$$

Considering the relations (3.30) and (3.31),

$$\begin{aligned}
J(\tilde{x}_{k+j}, \tilde{\mathbf{U}}_{k+j|k+j}) &\leq \sum_{i=0}^{N-1-j} d(x^*_{k+j+i|k}, \mathcal{T}_{x,j+i}, Q) + d(u^*_{k+j+i|k}, \mathcal{T}_{u,j+i}, Q) \\
&= \sum_{i=j}^{N-1} d(x^*_{k+i|k}, \mathcal{T}_{x,i}, Q) + d(u^*_{k+i|k}, \mathcal{T}_{u,i}, Q) \\
&\leq \sum_{i=j}^{N-1} d(x^*_{k+i|k}, \mathcal{T}_{x,i}, Q) + d(u^*_{k+i|k}, \mathcal{T}_{u,i}, Q) \\
&\quad + \underbrace{\sum_{i=0}^{j-1} d(x^*_{k+i|k}, \mathcal{T}_{x,i}, Q) + d(u^*_{k+i|k}, \mathcal{T}_{u,i}, R)}_{\neq 0 \text{ if } J(x^*_{k|k}, \mathbf{U}^*_{k|k}) \neq 0} \\
&= J(x^*_{k|k}, \mathbf{U}^*_{k|k}).
\end{aligned}$$

Thus far, we showed that $J(x_k, \mathbf{U}^*_{k|k}) - J(x_{k+j}, \tilde{\mathbf{U}}_{k+j|k+j}) \leq 0$ if $e_{k+j|k} \in \mathscr{E}_{j,k}$.

(*Triggering cost function decay*) In next step, we show that when a triggering is occurred, the cost function (3.9) decreases with respect to the previous triggering instant. To this end, let the triggering condition be enabled at $k+j+1$, i.e., either $e_{k+j+1|k} \notin \mathscr{E}_{j+1,k}$

for $j \in \mathsf{N}_{[N-2]}$ or $j = N-1$. Consider the sequences $\hat{\mathbf{U}}_{k+j+1|k+j+1}$ and $\hat{\mathbf{X}}_{k+j+1|k+j+1}$, given in (3.28), as the candidate solutions for the problem $\mathscr{P}(x_{k+j+1})$. In a nutshell, this part amounts to deriving a non-negative lower bound on $J(x_k, \mathbf{U}^*_{k|k}) - J(x_{k+j+1}, \mathbf{U}^*_{k+j+1|k+j+1})$ that decreases as $k \to \infty$. One can observe that, for all $i \in \mathsf{Z}_{[N-2]}$ and for all $w_k \in \mathscr{W}$:

$$\begin{aligned} d(\hat{u}_{k+j+1+i|k+j+1}, \mathcal{T}_{u,i}, R) &\leq d(\tilde{u}_{k+i+1|k+j}, \mathcal{T}_{u,i} \sim K_i L_i \mathscr{W}, R) \\ &= d(\tilde{u}_{k+i+1|k+j}, \mathcal{T}_{u,i+1}, R), \qquad (3.32a) \\ d(\hat{x}_{k+j+1+i|k+j+1}, \mathcal{T}_{x,i}, Q) &\leq d(\tilde{x}_{k+i+1|k+j}, \mathcal{T}_{x,i} \sim L_i \mathscr{W}, Q) \\ &= d(\tilde{x}_{k+i+1|k+j}, \mathcal{T}_{x,i+1}, Q), \qquad (3.32b) \end{aligned}$$

where the above inequalities are an immediate consequence of the definition (3.28) and Lemma 3.2.3. We also have $\hat{x}_{k+j+N|k+j+1} = \tilde{x}_{k+j+N|k+j} + L_{N-1} w_{k+j}$ and $\hat{u}_{k+j+N|k+j+1} = F\tilde{u}_{k+j+N|k+j+1} + FL_{N-1} w_{k+j}$ based on the definition (3.28). Recall that $L_{N-1} = 0_{n_x \times n_x}$ (from the nilpotency) and we have also shown that $\tilde{x}_{k+j+N|k+j} \in \mathscr{X}_f = \mathscr{R}$ in the proof of Theorem 3.4.3. It follows that $\hat{x}_{k+j+N|k+j+1} \in \mathcal{T}_{x,N-1}$ from (3.7c), and that $\hat{u}_{k+j+N|k+j+1} \in \mathcal{T}_{u,N-1}$ from (3.7e). Thus, the costs associated to $\hat{x}_{k+j+N|k+j+1}$ and $\hat{u}_{k+j+N|k+j+1}$ become zero for the candidate sequences. This last observation along with the inequalities given in (3.32) lead to

$$\begin{aligned} &J(x_{k+j+1}, \hat{\mathbf{U}}_{k+j+1|k+j+1}) = \\ &\sum_{i=0}^{N-1} d(\hat{x}_{k+j+1+i|k+j+1}, \mathcal{T}_{x,i}, Q) + d(\hat{u}_{k+j+1+i|k+j+1}, \mathcal{T}_{u,i}, R) \\ &\quad \leq \sum_{i=0}^{N-2} d(\tilde{x}_{k+j+1+i|k+j}, \mathcal{T}_{x,i+1}, Q) + d(\tilde{u}_{k+j+1+i|k+j}, \mathcal{T}_{u,i+1}, R) \\ &\quad = J(\tilde{x}_{k+j|k+j}, \tilde{\mathbf{U}}_{k+j|k+j}) - d(\tilde{x}_{k+j|k+j}, \mathcal{T}_{x,0}, Q) - d(\tilde{u}_{k+j|k+j}, \mathcal{T}_{u,0}, R). \end{aligned} \qquad (3.33)$$

As a result,

$$\begin{aligned} &J(x_{k+j+1}, \mathbf{U}^*_{k+j+1|k+j+1}) \\ &\quad \overset{\text{(i)}}{\leq} J(x_{k+j+1}, \hat{\mathbf{U}}_{k+j+1|k+j+1}) \\ &\quad \overset{\text{(ii)}}{\leq} J(\tilde{x}_{k+j|k+j}, \tilde{\mathbf{U}}_{k+j|k+j}) - d(\tilde{x}_{k+j|k+j}, \mathcal{T}_{x,0}, Q) - d(\tilde{u}_{k+j|k+j}, \mathcal{T}_{u,0}, R) \\ &\quad \overset{\text{(iii)}}{=} J(x_{k+j}, \tilde{\mathbf{U}}_{k+j|k+j}) - d(x_{k+j}, \mathcal{T}_{x,0}, Q) - d(u^*_{k+j|k+j}, \mathcal{T}_{u,0}, R) \\ &\quad \overset{\text{(iv)}}{\leq} J(x_k, \mathbf{U}^*_{k|k}) - d(x_{k+j}, \mathcal{T}_{x,0}, Q) - d(u^*_{k+j|k+j}, \mathcal{T}_{u,0}, R), \end{aligned}$$

where the inequality (i) follows from the optimality, the derivation (3.33) implies the inequality (ii), the equality (iii) follows from the facts that $\tilde{x}_{k+j|k+j} = x_{k+j}$ and $\tilde{u}_{k+j|k+j} = u^*_{k+j|k+j}$ (see the proof of Theorem 3.4.3), and finally, the conclusion made in the first part (*Inter-triggering cost function decay*) leads to the inequality (iv). Notice that the amount of reduction in the cost function J is $d(x_{k+j}, \mathcal{T}_{x,0}, Q) + d(u^*_{k+j|k+j}, \mathcal{T}_{u,0}, R)$. Since the cost function $J \geq 0$, the above analysis guarantees that $J(x_{k_{\text{trig}}}, \mathbf{U}^*_{k_{\text{trig}}|k_{\text{trig}}})$ converges to a steady value as the triggering instant $k_{\text{trig}} \to \infty$. This in turn implies $d(x_{k+j}, \mathcal{T}_{x,0}, Q) + d(u^*_{k+j|k+j}, \mathcal{T}_{u,0}, R) \to 0$, as $k \to \infty$, from which we can infer that $x_k \to \mathbb{T}_x$ and $u_k \to \mathbb{T}_u$ as $k \to \infty$. This concludes the proof.

3.5.5. PROOF OF THEOREMS 3.4.6 & 3.4.7

We first begin with a preliminary argument that is shared between both theorems. We then carry on with the proof of each case in an orderly fashion. Notice that $\xi \in \mathscr{S} \sim M\mathscr{B}(l,u)$ and $\mathscr{S}$ is a polytope by the theorems' hypothesis. By virtue of the relation (3.1), one can infer that

$$\langle a_{i,\mathscr{S}}^\top, \xi \rangle \leq b_{i,\mathscr{S}} - h_{M\mathscr{B}}(a_{i,\mathscr{S}}^\top), \text{ for all } i \in \mathsf{N}_{[m]}.$$

Next, observe that $\mathscr{B}(l,u) \subset \mathbb{R}^k$ is a polytope (and as a result bounded), the domain $\mathscr{K}_{\mathscr{B}}$ on which the support function $h_{\mathscr{B}}$ is defined is the whole space, i.e., $\mathscr{K}_{\mathscr{B}} = \mathbb{R}^k$. Hence, $h_{M\mathscr{B}}(a_{i,\mathscr{S}}^\top) = h_{\mathscr{B}}(M^\top a_{i,\mathscr{S}}^\top)$, and as a consequence

$$\langle a_{i,\mathscr{S}}^\top, \xi \rangle \leq b_{i,\mathscr{S}} - h_{\mathscr{B}}(M^\top a_{i,\mathscr{S}}^\top), \text{ for all } i \in \mathsf{N}_{[m]}.$$

Rearranging the above inequality, we arrive at

$$h_{\mathscr{B}}(M^\top a_{i,\mathscr{S}}^\top) \leq b_{i,\mathscr{S}} - \langle a_{i,\mathscr{S}}^\top, \xi \rangle, \text{ for all } i \in \mathsf{N}_{[m]},$$

where the only unknown entity is $h_{\mathscr{B}}(M^\top a_{i,\mathscr{S}}^\top)$ with $M^\top a_{i,\mathscr{S}}^\top \in \mathbb{R}^k$. It follows from the definition of the support function that $\langle M^\top a_{i,\mathscr{S}}^\top, z \rangle \leq h_{\mathscr{B}}(M^\top a_{i,\mathscr{S}}^\top)$ for all $z \in \mathbb{R}^k$. Thus,

$$\langle M^\top a_{i,\mathscr{S}}^\top, z \rangle \leq b_{i,\mathscr{S}} - \langle a_{i,\mathscr{S}}^\top, \xi \rangle, \text{ for all } i \in \mathsf{N}_{[m]}, \text{ for all } z \in \mathscr{B}. \tag{3.34}$$

Let us now define for all $i \in \mathsf{N}_{[m]}$

$$\begin{aligned} a_{i,\bar{\mathscr{S}}}^\top &:= M^\top a_{i,\mathscr{S}}^\top, \\ b_{i,\bar{\mathscr{S}}} &:= b_{i,\mathscr{S}} - \langle a_{i,\mathscr{S}}^\top, \xi \rangle, \end{aligned}$$

and the convex polytope (which we referred to as the *principal* polytope in the paragraph before Theorem 3.4.6)

$$\begin{aligned} \bar{\mathscr{S}} &:= \{ s \in \mathbb{R}^k : \langle a_{i,\bar{\mathscr{S}}}^\top, s \rangle \leq b_{i,\bar{\mathscr{S}}}, \text{ for all } i \in \mathsf{N}_{[m]} \} \\ &= \{ s \in \mathbb{R}^k : A_{\bar{\mathscr{S}}} s \leq b_{\bar{\mathscr{S}}} \}, \end{aligned} \tag{3.35}$$

where $A_{\bar{\mathscr{S}}} := [a_{1,\bar{\mathscr{S}}}^\top, \cdots, a_{m,\bar{\mathscr{S}}}^\top]^\top = (M^\top A_{\mathscr{S}}^\top)^\top = A_{\mathscr{S}} M$ and $b_{\bar{\mathscr{S}}} := [b_{1,\bar{\mathscr{S}}}, \cdots, b_{m,\bar{\mathscr{S}}}]^\top = b_{\mathscr{S}} - A_{\mathscr{S}} \xi$. Now, one can deduce from the inequalities (3.34) and the definition (3.35) that the convex polytope $\bar{\mathscr{S}}$ contains the hyper-rectangle $\mathscr{B}(l,u)$, i.e., $\mathscr{B}(l,u) \subseteq \bar{\mathscr{S}}$. Notice that $\mathscr{B}(l,u)$ is parametric in the variables l and u.

Theorem 3.4.6: In the CP framework, we propose a convex nonlinear program to compute the hyper-rectangle $\mathscr{B}(l,u) \subseteq \bar{\mathscr{S}}$ such that its volume is maximized. Suppose $\mathscr{B}(l,u)$ is parameterized as $l := -\underline{v} = [-\underline{v}_1, \cdots, -\underline{v}_k]^\top$ and $u := \overline{v} = [\overline{v}_1, \cdots, \overline{v}_k]^\top$ such that for all $i \in \mathsf{N}_{[k]}$, $\underline{v}_i$ and $\overline{v}_i$ are positive scalars (the positiveness condition has to do with the fact that the resulting hyper-rectangle should contain the origin). Recall the inequality (3.34), that is $\langle M^\top a_{i,\mathscr{S}}^\top, z \rangle \leq b_{i,\mathscr{S}} - a_{i,\mathscr{S}} \xi$, for all $i \in \mathsf{N}_{[m]}$, for all $z \in \mathscr{B}$. In what follows, we show that although the hyper-rectangle $\mathscr{B}(l,u) = \mathscr{B}(-\underline{v}, \overline{v})$ is parametric, one can

provide a closed-form for its support function evaluated at $M^\top a_{i,\mathscr{S}}^\top$. By definition of a support function,

$$\begin{aligned} h_{\mathscr{B}}(M^\top a_{i,\mathscr{S}}^\top) = \max_z \ & \langle M^\top a_{i,\mathscr{S}}^\top, z\rangle \\ \text{s.t. } & A_{\mathscr{B}} z \le b_{\mathscr{B}}, \end{aligned} \tag{3.36}$$

where $A_{\mathscr{B}} = [\mathsf{I}_k \ -\mathsf{I}_k]^\top$ and $b_{\mathscr{B}} = [\overline{v}^\top \ \underline{v}^\top]^\top$. The above problem is an LP with a bounded feasible set. Thus, the optimal solution lies on the boundary of the hyper-rectangle to which the normal $M^\top a_{i,\mathscr{S}}^\top$ points toward. Let us define, for all $i \in [m]$, $\hat{w}^i := \text{sign}(M^\top a_{i,\mathscr{S}}^\top)$, where the sign operator is applied entry-wise. (Notice that this vector simply shows to which orthant(s) the vector $M^\top a_{i,\mathscr{S}}^\top$ points to and $\hat{w}^i \in \mathbb{R}^k$.) It then becomes clear the vectors $w^i \in \mathbb{R}^{2k}$, as defined in (3.21), enable us to express the optimal solution of (3.36) in terms of a linear combination of the vertices of $\mathscr{B}$, i.e.,

$$h_{\mathscr{B}}\big(M^\top a_{i,\mathscr{S}}^\top\big) = \langle w^i, [\overline{v}^\top \ \underline{v}^\top]^\top\rangle, \text{ for all } i \in \mathsf{N}_{[m]}.$$

Based on the above relation, the inequality (3.34) simplifies to

$$\langle w^i, [\overline{v}^\top \ \underline{v}^\top]^\top\rangle \le b_{i,\mathscr{S}} - a_{i,\mathscr{S}}\xi, \text{ for all } i \in \mathsf{N}_{[m]},$$

in which the vectors $\underline{v}, \overline{v} \in \mathbb{R}^k$ are the decision variables. Intuitively, the above inequalities represents the linear constraints that the vertices of the hyper-rectangle $\mathscr{B}(-\underline{v}, \overline{v})$ should satisfy in order to guarantee $\xi \in \mathscr{S} \sim M\mathscr{B}(-\underline{v}, \overline{v})$.

Based on the chosen definition of volume for $\mathscr{B}(-\underline{v}, \overline{v})$ in (3.17), we intend to find a hyper-rectangle $\mathscr{B}(-\underline{v}, \overline{v})$ that possesses the maximal volume. Unfortunately, regardless of the definition choice for the volume, the resulting objective function is non-convex and becomes unsuitable for optimization. Interestingly enough, one can simply use the logarithmic mapping for the volume definitions in (3.17) to obtain the objective functions suggested in (3.20), that are monotonic nonlinear concave functions. Then, it follows that a maximum hyper-rectangle $\mathscr{B}$ that contains the origin and satisfies $\xi \in \mathscr{S} \sim M\mathscr{B}$ is the solution of the CP (3.19).

Theorem 3.4.7: In the LP framework, we follow the procedure proposed in [77] with which one is able to cast the problem as a linear program. We first provide the proof for the LP relaxation of the problem (3.19) with $q = 1$. Let us denote the maximum length of a line segment containing the origin, parallel to the j-th coordinate axis, and contained in $\bar{\mathscr{S}}$ by r_j. It follows from [77, Proposition3] that one can use (3.23) to find r_j, for all $j \in \mathsf{N}_{[k]}$. It is worth nothing that in the LP (3.23), the constraints $z \le 0$ and $z + \omega e_j \ge 0$ are two extra regularity conditions that we placed on the line segment compared to [77, Proposition3]. These conditions ensure that the origin lies inside this line segment. Now, define the strictly positive vector $r \in \mathbb{R}^k$ by $r_j = \omega_j$ for all $j \in \mathsf{N}_{[k]}$. Then, it follows from [77, Proposition2] that a maximum r-constrained inner hyper-rectangular $\mathscr{B}$ of $\bar{\mathscr{S}}$ that contains the origin is given by $\mathscr{B}(z^*, z^* + \lambda^* r)$ where z^* and λ^* are the optimal solutions of (3.22). Here, we also emphasize the fact that we have introduced the extra constraints $z \le 0$ and $z + \lambda r \ge 0$ with respect to [77, Proposition2]. By doing so, the LP (3.22) is forced to find a hyper-rectangular $\mathscr{B}$ such that it contains the origin. Then, the claim for the LP case follows.

We now present a sketch of proof for the LP relaxation of the problem (3.19) with $q = 2$. Observe that the polytope $\bar{\mathscr{P}}' := \{s \in \mathbb{R}^{2k} : W's \leq B'\}$ is the matrix representation of the constraints in the CP (3.19), where W' and B' are defined in Theorem 3.4.7. We seek to find a hyper-rectangle that fits inside this *lifted* polytope as follows. In the first step, we place a vertex of the hyper-rectangle at the origin. We then find the width of the line segment along each coordinate that is inside the lifted polytope and contains the origin using (3.25). In the second step, we use (3.24) to find a scaling factor λ such that the λ-scaled hyper-rectangle constructed based on the first step fits inside the polytope $\bar{\mathscr{P}}'$. This concludes the proof.

3.6. NUMERICAL EXAMPLES

In this section, we provide a numerical example to study the results presented in Section 3.4. Consider the perturbed LTI system

$$x_{k+1} = \begin{pmatrix} 1 & 0.1 \\ 0 & 1 \end{pmatrix} x_k + \begin{pmatrix} 0.005 \\ 0.1 \end{pmatrix} u_k + w_k,$$

where the states and input constraint sets are $\mathbb{X} = \{x \in \mathbb{R}^2 : |x_1| \leq 5,\ |x_2| \leq 3\}$ and $\mathbb{U} = \{u \in \mathbb{R} : |u| \leq 5\}$, respectively. The disturbance set is defined as $\mathscr{W} = \{w \in \mathbb{R}^2 : w_1 = 0, |w_2| \leq 0.1\}$. The state and input target sets are $\mathbb{T}_x = \{x \in \mathbb{R}^2 : \|x\|_\infty \leq 1\}$ and $\mathbb{T}_u = \{u \in \mathbb{R} : |u| \leq 3\}$, respectively. The horizon length N in the optimization problem (3.10) is set to 25. The weight matrices in the cost function (3.9) are

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & 0.001 \end{pmatrix} \text{ and } R = 10.$$

Finally, the terminal set $\mathscr{X}_f = \{x \in \mathbb{R}^2 : \|x\|_\infty \leq 0.1\}$.

At some triggering instant $k \in \mathbb{Z}_{\geq 0}$, let us first compare the four construction frameworks for the sequence of hyper-rectangles $\{\mathscr{E}_{j,k}\}_{j=1}^{N-1}$ provided in Theorems 3.4.6 & 3.4.7. Figure 3.2 depicts a subset of the constructed hyper-rectangles $\{\mathscr{E}_{j,k}\}_{j=1}^{N-1}$. We now argue in the support of Remark 3.4.9 regarding the directional sensitivity of the constructed hyper-rectangles and the fact that using the volume definition (3.17b) instead of the standard definition of the volume (3.17a) will reduce the level of directional sensitivity in the constructed hyper-rectangles. For example, consider the instant $k+12$ (i.e., $j = 12$). The (green) filled polytope constructed around $x^*_{k+12|k}$ represents the principal polytope $\bar{\mathscr{P}}_{12,k}$, i.e., the polytope that represents the constraints (3.18b)-(3.18f). Moreover, the hyper-rectangles $\mathscr{E}_{12,k}$ constructed around $x^*_{k+12|k}$ using CP_1, LP_1, CP_2, and LP_2 are depicted in (green) dash-dotted, dashed, dotted, solid rectangles, respectively. The constructed hyper-rectangles based on CP_1 and LP_1 are extremely sensitive to the direction of the prediction error $e_{k+12|k} = x_{k+12} - x^*_{k+12|k}$ along the x_1 and x_2 axes. On the other hand, it is evident that the level of directional sensitivity is reduced in the hyper-rectangles that are constructed based on CP_2 and LP_2. Another matter that should be highlighted in Figure 3.2 is as follows. As the inter-event instant $j \to N-1$ ($j \in \mathsf{N}_{[N-1]}$), the principal polytopes $\bar{\mathscr{P}}_{j,k}$ (and as a result, the constructed hyper-rectangles $\mathscr{E}_{j,k}$) become more symmetric with respect to $x^*_{k+j|k}$. In terms of Remark 3.4.9, the constructed hyper-rectangles $\mathscr{E}_{j,k}$ thus become less directionally sensitive as $j \to N-1$.

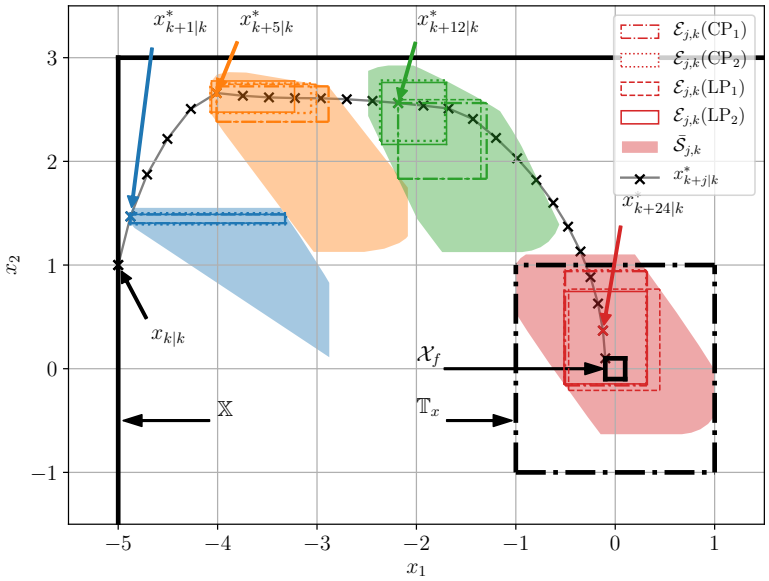


Figure 3.2: Comparison of the CP and LP frameworks of Theorems 3.4.6 & 3.4.7 to construct the hyper-rectangles $\{\mathscr{E}_{j,k}\}_{j=1}^{N-1}$. The optimal state trajectory $\mathbf{X}^*_{k|k} = \{x^*_{k+i|k}\}_{i=0}^{N}$ is depicted by the (gray) solid/cross line. For several instant over the horizon $N = 25$, the (colored) dash-dotted, dashed, dotted, solid rectangles represent the hyper-rectangles $\mathscr{E}_{j,k}$ using CP_1, LP_1, CP_2, and LP_2, respectively. For some $j \in \mathsf{N}_{[N-1]}$, the (colored) filled polytopes constructed around $x^*_{k+i|k}$ represent the principal polytopes $\bar{\mathscr{S}}_{j,k}$.

We next consider two types of disturbance realizations by employing the results of Theorems 3.4.6 & 3.4.7. We also use Algorithm 2 to control the considered perturbed LTI system. The first case deals with a uniform disturbance ($|w_2| \leq 0.1$), see Figure 3.3, while the second case assumes a worst-case disturbance ($w_2 \in \{\pm 0.1\}$), see Figure 3.4.

First of all, it is evident that the input and state trajectories do not violate the constraint sets $\mathbb{T}_x$ and $\mathbb{T}_u$, respectively, in the both cases of the disturbance realizations. Moreover, the state x_k and the input u_k converge to the target sets $\mathbb{T}_x$ and $\mathbb{T}_u$, respectively. Nonetheless, the standard implementation in both cases of disturbance realizations converges to a smaller subset of the target set $\mathbb{T}_x$ compared to the four event-based implementations. Furthermore, as it has been claimed, the event-based implementations reach the target set $\mathbb{T}_x$ with a smaller number of triggering instants compared to the time-triggered standard implementation: the CP_1, LP_1, CP_2, LP_2, and standard implementations require 13, 13, 6, 6, and 21 number of triggering instants, respectively. However, in the case of worst-case disturbance, the number of triggering instants required is almost the same in all four cases. It is also interesting to note that the event-based implementations exhibit an almost limit-cyclic behavior inside the target set $\mathbb{T}_x$ in the worst-case disturbance realization.

3.7. CONCLUSIONS

In this chapter, we proposed an approach to apply an RMPC method to constrained, perturbed LTI systems in an event-based fashion. The procedure to design the triggering mechanism is decoupled from the controller design. Furthermore, we introduced two theoretical frameworks to construct the triggering mechanism as a volume maximization problem. One framework is a general nonlinear convex program while the other framework is a linear program. In particular, we proposed a non-standard definition of volume to address the limitations that occur in the case of using the standard definition of the volume in the process of designing the triggering mechanism. For each choice of the volume definition, our numerical experiments showed that the theoretical frameworks provide a similar behavior at the price of the convex program framework being more computationally expensive compared to the linear program framework. On the other hand, the linear and convex program frameworks based on the non-standard definition of the volume outperform the linear and convex program frameworks based on the standard definition of the volume. There are multiple directions that one can pursue to extend the results in this chapter. First, it is interesting to investigate the impact of disturbance gains $\{K_i\}_{i=0}^{N-1}$ on the sparsity pattern of the convex programs in Theorems 3.4.6 & 3.4.7. By doing so, one can customize the solver to speed up the triggering mechanism's design. Second, we have shown in the proof of Theorem 3.4.4 that the performance cost per average transmission in the event-triggered implementation is smaller than the one in the standard implementation (notice that our setting is deterministic). It would be interesting to investigate how this observation is related to the few studies in the literature (e.g., [65] which is in a stochastic setting) that guarantee superiority of an event-based implementation with respect to a standard implementation.

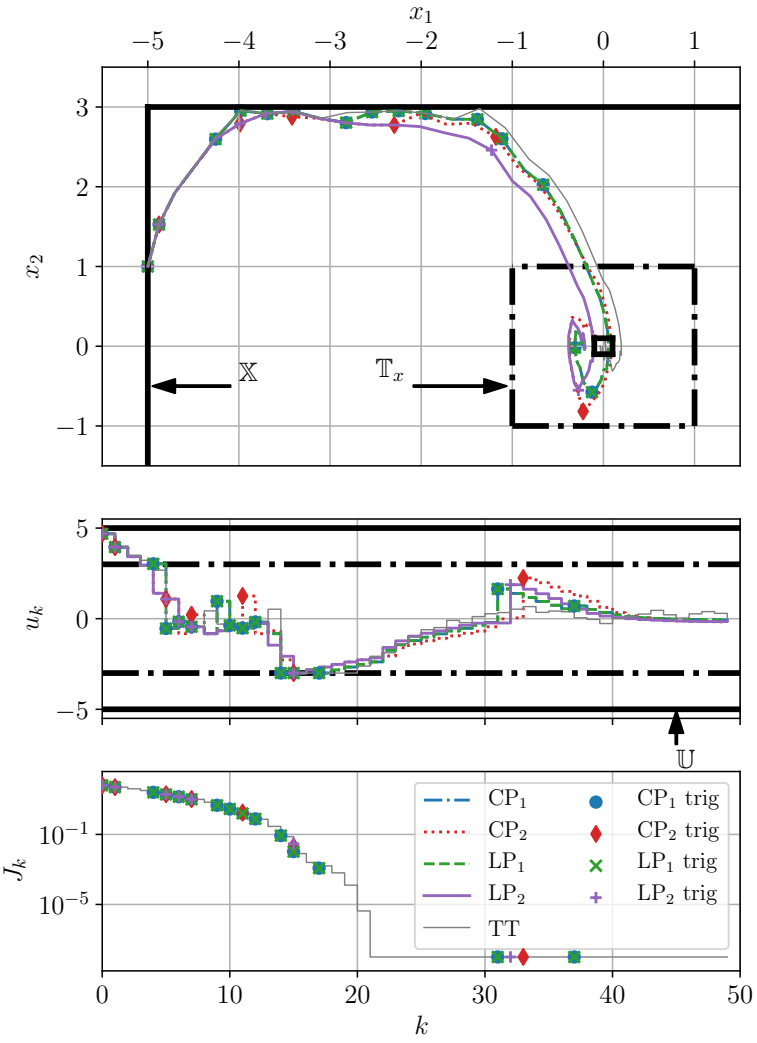


Figure 3.3: (Uniform case disturbance) Comparison of the CP and LP construction approaches of the event-based implementation along with the standard time-triggered implementation. The (blue) dash-dotted/circle, (green) dashed/cross, (red) dotted/diamond, (purple) solid/plus, and (gray) thin lines are associated with the event-based implementation using the CP$_1$, LP$_1$, CP$_2$, and LP$_2$ construction methods, and the standard time-triggered implementation, respectively. (Top) Phase portrait. (Middle) Input sequence. (Bottom) Cost function.

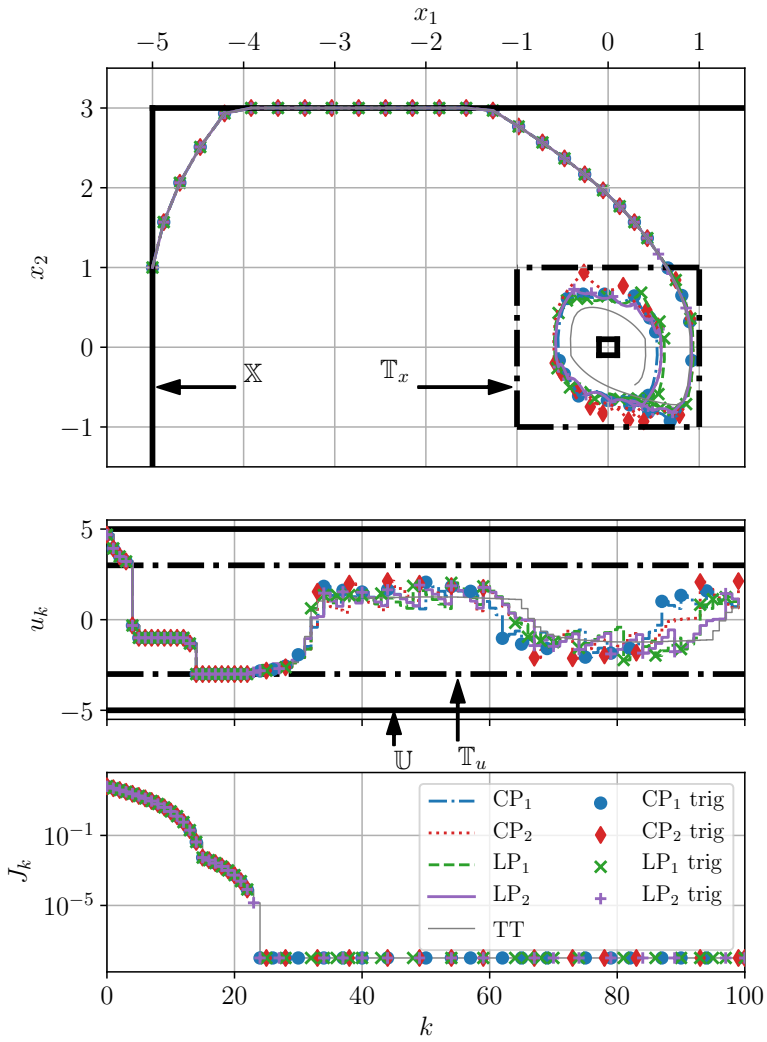


Figure 3.4: (Worst case disturbance) Comparison of the CP and LP construction approaches of the event-based implementation along with the standard time-triggered implementation. The (blue) dash-dotted/circle, (green) dashed/cross, (red) dotted/diamond, (purple) solid/plus, and (gray) thin lines are associated with the event-based implementation using the CP_1, LP_1, CP_2, and LP_2 construction methods, and the standard time-triggered implementation, respectively. (Top) Phase portrait. (Middle) Input sequence. (Bottom) Cost function.

4

TIMING ABSTRACTION OF AN EVENT-TRIGGERING MECHANISM

In networked control systems, the introduction of event-triggering strategies in the sampling process has led to possible usage reductions in certain network capacities, such as the communication bandwidth. However, these possible beneficiary properties come at a price. Due to the aperiodic nature of sampling periods generated by these strategies, the schedulability problem becomes much more demanding compared to the traditional, periodic strategy. This chapter is an attempt to address this issue by using tools from formal verification. We focus on perturbed linear time-invariant systems with an $\mathscr{L}_2$-based triggering mechanism. Inspired by an approach in the literature, we introduce a framework to construct a timed safety automaton that captures the aperiodic sampling behavior of the considered class of control systems. In this framework, the state-space is partitioned into a finite number of convex polyhedral cones, each of which represents a discrete mode in the abstracted automaton. Adopting techniques from stability analysis of retarded systems accompanied with a polytopic embedding of time, LMI conditions to characterize the sampling interval associated with each cone are derived. We then use some tools from reachability analysis of linear systems to derive all the transitions in the abstracted automaton. The materials presented in this chapter are previously reported in [6].

4.1. INTRODUCTION

Wireless networked controlled systems (WNCSs) represent a class of spatially distributed control systems in which feedback loops are closed via shared communication components. Several advantages of WNCSs, such as the ease of maintenance and the flexibility of implementation, make them an attractive choice in industrial environments. Meanwhile, WNCSs are burdened with shortcomings, such as limited battery life and communication bandwidth. Under these circumstances, the resource over-utilization caused by (traditional) periodic implementations, the so-called time-driven control (TDC), makes such implementations less appealing for WNCSs.

To address the above issues, control researchers have proposed *event-driven control* (EDC) strategies that are aperiodic, such as *event-triggering control* (ETC) [47] and *self-triggering control* (STC) [81]. In EDC strategies, the core idea is as follows. The dynamics of the control system determine the next sampling instant in the hope of attenuating the usage of resources. In these strategies, control task executions only happen when a pre-specified condition is violated. This condition is called the *triggering mechanism* (TM). It is derived based on stability and/or performance of the closed-loop system. On the other hand, the *schedulability* of ETC strategies, due to their aperiodic nature, is more arduous compared to TDC strategies. In fact, in TDC strategies, the control and scheduler designs are naturally decoupled via the (pre-defined) fixed sampling period. This phenomenon is called the *separation-of-concerns* in the real-time systems community [82]. It is worth mentioning that ETC strategies are almost always equipped with a *minimum inter-execution time* (MIET) to prevent the occurrence of *Zeno* behavior in the sampling process. This quantity can be technically used in the synthesis of a task scheduler. However, it is a conservative approximation of the lower bound on all the possible generated sampling periods. Thus, such synthesis does not make use of the beneficiary characteristics of ETC strategies in an efficient manner. To address this shortcoming, researchers have proposed another class of approaches, the so-called *co-design* approaches. In this

class, the problem of controller and scheduler syntheses for real-time systems are tackled in a unified framework, see e.g., feedback modification to task attributes [83], [84], [85], [86], anytime controllers [87], [88], and event-based control and scheduling [89], [90].

Recently, alternative to the unified frameworks mentioned above, [91], [92] have proposed a decoupling framework to capture the sampling behavior of linear time-invariant (LTI) systems with ISS-based TMs using timed safety automata (TSAs). Generally speaking, TSA is a simplified version of timed automaton (TA) [93], [94]. It is a powerful tool to model the timing behavior of real-time systems for scheduling purposes since its reachability analysis is decidable [95], [96]. In this chpater, following the same path as in [91], [92], we propose a framework to capture the sampling behavior of perturbed LTI systems with the $\mathscr{L}_2$-based TM proposed by [97]. We show that the derived TSA ε-approximately simulates the sampling behavior of the $\mathscr{L}_2$-based ETC system. It is evident that such characterizations can be analyzed independently from a scheduling perspective, thus providing a scalable and versatile framework to design event-triggering WNCSs.

The rest of this chapter is organized as follows. The preliminary notions are introduced in Section 4.2. The abstraction of the considered ETC approach is presented in Section 4.3. Section 4.4 provides the technical proofs of the main results. A numerical example to validate the results of this chapter is given in Section 4.5. The chapter is concluded in Section 4.6.

Notations: $\mathbb{R}^n$ and $\mathbb{R}_{>0}$ denote the n-dimensional Euclidean space and the positive reals, respectively. $\mathbb{N}$ is the set of positive integers, $\mathbb{R}_{\geq 0}$ represents the set of nonnegative reals, and $\mathbb{IR}^+$ is the set of all closed intervals $[a,b]$ such that $a,b \in \mathbb{R}_{>0}$ and $a \leq b$. For any set S, 2^S denotes the set of all subsets of S, i.e. the power set of S. $\mathscr{S}_{m\times n}$ and $\mathscr{S}_n$ are the set of all $m \times n$ real-valued matrices and the set of all $n \times n$ real-valued symmetric matrices, respectively. For a matrix M, $M \preceq 0$ (or $M \succeq 0$) means M is a negative (or positive) semidefinite matrix and $M \prec 0$ ($M \succ 0$) indicates M is a negative (positive) definite matrix. Also, $M^\top$ is the transpose of M. $\mathscr{S}_n^+$ is the cone of all $n \times n$ symmetric positive definite matrices. $\lfloor x \rfloor$ indicates the largest integer not greater than $x \in \mathbb{R}$. $|y|$ and $\|M\|$ denote the Euclidean norm of a vector $y \in \mathbb{R}^n$ and the Frobenius norm of a matrix $M \in \mathscr{S}_{m\times n}$, respectively. For a matrix $M \in \mathscr{S}_n$, $\lambda(M)$ and $\lambda_{\max}(M)$ denote the set of eigenvalues and the largest eigenvalue of M. Consider two sets $X, Y \subseteq \mathbb{R}^n$, their Minkowski sum is given by $X \oplus Y := \{x+y | x \in X \text{ and } y \in Y\}$.

4.2. PRELIMINARIES

We recall some notions that are employed in the remainder of this chapter. We begin with collecting some results from convex analysis. We next describe the considered ETC system. Followed by that, a concise review of some notions from formal verification is provided. Then, the notion of TSA as a tool that can capture the triggering behavior of ETC systems is presented. Finally, the considered problem addressed in this chapter is formally stated.

We state the following known results that will be used in Subsection 4.2.1.

Lemma 4.2.1. *([98, Lemma 6.2]) For any real matrices E, G and real symmetric positive*

definite matrix P, with compatible dimensions,

$$EG + G^\top E^\top \preceq EPE^\top + G^\top P^{-1} G.$$

Lemma 4.2.2. *([99, Section 2]) For all $A \in \mathscr{S}_{n\times n}$, if $\mu(A) = \max\{\mu \in \mathbb{R} |\ \mu \in \lambda\left(\frac{A^\top + A}{2}\right)\}$, then, $|e^{At}| \leq e^{\mu(A)t}$.*

Proposition 4.2.3. *(Jensen Inequality [98, Proposition B.8]) For any matrix $M \in \mathscr{S}_m^+$ with constant entries, scalar $\gamma > 0$, vector function $\omega : [0,\gamma] \to \mathbb{R}^m$ such that the integrations concerned are well defined, then,*

$$\gamma \int_0^\gamma \omega^\top(\beta) M \omega(\beta) d\beta \geq \left(\int_0^\gamma \omega(\beta) d\beta\right)^\top M \int_0^\gamma \omega(\beta) d\beta.$$

4

4.2.1. $\mathscr{L}_2$-BASED ETC SYSTEM

In this subsection, an overview of the ETC strategy proposed by [97] along with a new result (see Theorem 4.2.4) are presented. Consider a sampled-data system

$$\begin{aligned} \dot{\xi}(t) &= A\xi(t) + B\nu(t) + E\omega(t), \forall\, t \in [0, \tau(x)), \\ \xi(0) &= x, \end{aligned} \tag{4.1}$$

where $\xi(t) \in \mathbb{R}^n$, $\nu(t) \in \mathbb{R}^m$, $\omega(t) \in \mathbb{R}^p$, $\tau(x)$ denotes the sampling period associated with $\xi(0)$, and the matrices A, B, and E have compatible dimensions. The control law

$$\nu(t) = -Kx, \tag{4.2}$$

is implemented in a sample-and-hold manner. Furthermore, assume that the disturbance ω is a vanishing type disturbance [97], i.e.,

$$\text{there exists } W \geq 0 \text{ such that } |w(t)|^2 \leq W|x|^2, \text{ for all } t \in [0, \tau(x)). \tag{4.3}$$

Denote by ϵ, the error signal endured by the system (4.1)-(4.2), $\epsilon(t) = x - \xi_x(t)$ where $\xi_x(t)$ is the solution of (4.1). Reformulating (4.1), the evolution of state and error signals can be rewritten in the compact form

$$\xi_x(t) = \Lambda(t)x + \Omega(t), \tag{4.4a}$$

and

$$\epsilon(t) = \big(I - \Lambda(t)\big)x - \Omega(t), \tag{4.4b}$$

where

$$\Lambda(t) = I + \int_0^t e^{As} ds (A - BK), \qquad \Omega(t) = \int_0^t e^{A(t-s)} E\omega(s) ds. \tag{4.5}$$

Assume that there exists a quadratic Lyapunov function $V(\xi) = \xi^\top P \xi$ such that P is the solution to the Algebraic Riccati Equation (ARE)

$$PA + A^\top P - Q + R = 0, \tag{4.6a}$$

where

$$Q = PBB^\top P, \quad R = \frac{1}{\gamma^2} PEE^\top P, \quad \gamma > 0. \tag{4.6b}$$

Notice that the existence of V guarantees that the system (4.1) with the full-state feedback $v(t) = -K\xi(t) = -B^\top P\xi(t)$ is finite-gain $\mathscr{L}_2$ stable from ω to $(x^\top, u^\top)$ with an induced gain less than γ [97].

The authors of [97] proposed the state-dependent TM

$$\tau(x) := \inf\{t > 0 : \epsilon^\top(t) M\epsilon(t) \geq x^\top N x\}, \tag{4.7}$$

where

$$M = (1-\beta^2)I + PBB^\top P, \quad N = \frac{1}{2}(1-\beta^2)I + PBB^\top P,$$

and $\beta > 0$ is a user-defined scalar related to the TM (4.7).

In what follows, we provide a certain type of LMI conditions with which a lower bound on the inter-sample time can be derived.

Theorem 4.2.4 (Inter-sample lower bound)**.** *Consider the system (4.1)-(4.2) with the triggering mechanism (4.7). Assume that there exist a scalar μ and a symmetric matrix Ψ such that*

$$\mu \geq 0, \quad \Psi > 0, \quad M + \Psi \preceq \mu I, \tag{4.8a}$$

$$\Phi(t) \geq 0, \tag{4.8b}$$

where

$$\Phi(t) = \begin{bmatrix} \Phi_1(t) & \Phi_2(t) \\ \Phi_3(t) & \Phi_4(t) \end{bmatrix}, \tag{4.8c}$$

$$\begin{aligned} \Phi_1(t) &= \big(\Lambda(t) - I\big)^\top M\big(\Lambda(t) - I\big) + tW\mu\lambda_{\max}\big(E^\top E\big) d_A(t) I - N, \\ \Phi_2(t) &= \Phi_3^\top(t) = \big(\Lambda(t) - I\big)^\top M^\top, \\ \Phi_4(t) &= -\Psi, \end{aligned}$$

are satisfied. Then, the sampling period $\tau(x)$ generated by (4.7) is lower bounded by

$$\tau'(x) := \inf\{t > 0 : \Phi(t) \geq 0\}. \tag{4.9}$$

Proof. Let us first rewrite the TM (4.7) into a more suitable from. Substitute (4.4b) into (4.7). The TM (4.7) can be reformulated into

$$\tau(x) = \inf\{t > 0 : \mathscr{F}_\omega(x,t) \geq 0\}, \text{ for all } x \in \mathbb{R}^n, \tag{4.10}$$

where

$$\begin{aligned} \mathscr{F}_\omega(x,t) = x^\top &\Big[\big(\Lambda(t) - I\big)^\top M\big(\Lambda(t) - I\big) - N\Big] x \\ &+ x^\top \big(\Lambda(t) - I\big)^\top M\Omega(t) + \Omega^\top(t) M\big(\Lambda(t) - I\big) x + \Omega^\top(t) M\Omega(t), \end{aligned}$$

with $\Lambda(t)$ and $\Omega(t)$ as defined in (4.5). Let $\lambda^A_{\max}$ denote $\lambda_{\max}(A+A^\top)$ for the sake of compactness. Consider the terms in $\mathscr{F}_\omega(x,t)$ that are dependent on both of x and $\Omega(t)$. We employ Lemma 4.2.1 to decouple these terms, i.e.,

$$x^\top\big(\Lambda(t)-I\big)^\top M\Omega(t)+\Omega^\top(t)M\big(\Lambda(t)-I\big)x \leq \\ \Omega^\top(t)\Psi\Omega(t)+x^\top\big(\Lambda(t)-I\big)^\top M\Psi^{-1}M\big(\Lambda(t)-I\big)x,$$

for any matrix $\Psi=\Psi^\top\succ 0$. Hence,

$$\mathscr{F}_\omega(x,t)\leq x^\top\Big[\big(\Lambda(t)-I\big)^\top\big(M+M\Psi^{-1}M\big)\big(\Lambda(t)-I\big)-N\Big]x+\Omega^\top(t)(M+\Psi)\Omega(t).$$

Observe that

$$\begin{aligned}\Omega^\top(t)(M+\Psi)\Omega(t) &\leq \mu\Big(\int_0^t e^{A(t-s)}E\omega(s)ds\Big)^\top\Big(\int_0^t e^{A(t-s)}E\omega(s)ds\Big)\\ &\qquad\text{(assuming } M+\Psi\preceq\mu I \text{ and } \mu\geq 0)\\ &\leq t\mu\int_0^t e^{(t-s)\lambda^A_{\max}}\omega^\top(s)E^\top E\omega(s)ds\\ &\qquad\text{(using Lemma 4.2.2 and Propostion 4.2.3)}\\ &\leq tW\mu\lambda_{\max}\big(E^\top E\big)\Big(\int_0^t e^{\lambda^A_{\max}(t-s)}ds\Big)|x|^2\\ &\qquad\text{(using (4.3))}\\ &= tW\mu\lambda_{\max}\big(E^\top E\big)d_A(t)x^\top x,\end{aligned}$$

where

$$d_A(t)=\begin{cases}\dfrac{1}{\lambda^A_{\max}}\big(e^{\lambda^A_{\max}t}-1\big), & \lambda^A_{\max}\neq 0,\\ t, & \lambda^A_{\max}=0.\end{cases}$$

Based on the aforementioned procedure, one concludes

$$\mathscr{F}_\omega(x,t)\leq x^\top\Theta(t)x, \tag{4.11}$$

where

$$\Theta(t)=\big(\Lambda(t)-I\big)^\top\big(M+M\Psi^{-1}M\big)\big(\Lambda(t)-I\big)+tW\mu\lambda_{\max}\big(E^\top E\big)d_A(t)I-N. \tag{4.12}$$

Then, we employ the Schur complement in order to transform (4.12) into (4.8c). (Notice that the matrix Ψ appears in a nonlinear manner in the equality (4.12) while the equality (4.8c) is linearly dependent on Ψ.) Considering the inequality (4.11) and the fact that $\Phi(t)\geq 0$ implies $x^\top\Theta(t)x\geq 0$, it follows that $\tau(x)\geq\tau'(x)$ where τ and τ' are defined in (4.10) and (4.9), respectively. This concludes the proof. □

Remark 4.2.5 (Role of Theorem 4.2.4)**.** *Theorem 4.2.4 helps us to conservatively simplify the impacts of the unknown perturbation $\omega(t)$ in studying the behavior of the sampling function* (4.7)*. However, it is still intractable to employ the definition* (4.9) *in order to study the sampling behavior since this definition has to be checked for an infinite number of instants t. In addition, this definition lacks any insight on how the state x at the sampling instant affects the sampling period $\tau(x)$.*

4.2.2. SYSTEMS AND RELATIONS

In what follows, we review some notions from the field of system theory to formally characterize the outcome of our framework. Let Z be a set and $Q \subseteq Z \times Z$ be an equivalence relation on Z. Then, $[z]$ denotes the equivalence class of $z \in Z$ and Z/Q denotes the set of all equivalence classes. A metric (or simply a distance function) $d : Z \times Z \to \mathbb{R} \cup \{+\infty\}$ on Z satisfies, for all $x, y, z \in Z$: (i) $d(x,y) = d(y,x)$, (ii) $d(x,y) = 0 \leftrightarrow x = y$, and (iii) $d(x,y) \leq d(x,z) + d(y,z)$. The ordered pair (Z,d) is said to be a metric space.

Definition 4.2.6. *(Hausdorff Distance [100]) Assume X and Y are two non-empty subsets of a metric space (Z,d). The Hausdorff distance $d(X,Y)$ is*

$$\max\left\{\sup_{x\in X}\inf_{y\in Y} d(x,y), \sup_{y\in Y}\inf_{x\in X} d(x,y)\right\}.$$

4

It follows that the ordered pair $(\mathbb{IR}^+, d_H)$ is a metric space. We next introduce some concepts from system theory, in particular a modified notion of *quotient system* adopted from [91], see e.g., [101] for the traditional definition.

Definition 4.2.7 (System [101])**.** *A system is a sextuple $(X, X_0, U, \longrightarrow, Y, H)$ consisting of*

- *a set of states X,*
- *a set of initial states $X_0 \subseteq X$,*
- *a set of inputs U,*
- *a transition relation $\longrightarrow \subseteq X \times U \times X$,*
- *a set of outputs Y,*
- *an output map $H : X \to Y$.*

When the set of outputs Y of a system is endowed with a metric, it is called a *metric* system. An *autonomous* system is a system for which the cardinality of its input set is at most one.

Definition 4.2.8 (Approximate Simulation Relation [101])**.** *Consider two metric systems $S_a = (X_a, X_{a0}, U_a, \xrightarrow[a]{}, Y_a, H_a)$ and $S_b = (X_b, X_{b0}, U_b, \xrightarrow[b]{}, Y_b, H_b)$ with $Y_a = Y_b$, and let $\varepsilon \in \mathbb{R}_{\geq 0}$. A relation $R \subseteq X_a \times X_b$ is an ε-approximate simulation relation from S_a to S_b if the following conditions*

1. *for all $x_{a0} \in X_{a0}$, there exists $x_{b0} \in X_{b0}$ such that $(x_{a0}, x_{b0}) \in R$,*
2. *for all $(x_a, x_b) \in R$, we have $d(H_a(x_a), H_b(x_b)) \leq \varepsilon$,*
3. *for all $(x_a, x_b) \in R$, $(x_a, u_a, x'_a) \in \xrightarrow[a]{}$ in S_a, there exits $(x_b, u_b, x'_b) \in \xrightarrow[b]{}$ in S_b satisfying $(x'_a, x'_b) \in R$,*

are satisfied. We say that S_b ε-approximately simulates S_a, denoted by $S_a \preceq^{\varepsilon}_{\mathscr{S}} S_b$, if there exists an ε-approximate simulation relation R from S_a to S_b.

Definition 4.2.9 (Power Quotient System [91])**.** *Let $S = (X, X_0, \varnothing, \longrightarrow, Y, H)$ be an autonomous system and R be an equivalence relation on X. The power quotient of S by R, denoted by $S_{/R}$, is the autonomous system $(X_{/R}, X_{/R,0}, \varnothing, \xrightarrow[/R]{}, Y_{/R}, H_{/R})$ consisting of*

- $X_{/R} = X/R$,
- $X_{/R,0} = \{x_{/R} \in X_{/R} | x_{/R} \cap X_0 \neq \varnothing\}$,
- $(x_{/R}, u, x'_{/R}) \in \xrightarrow[/R]{}$ *if there exists* $(x, u, x') \in \longrightarrow$ *with* $x \in x_{/R}$ *and* $x' \in x'_{/R}$,
- $Y_{/R} \subset 2^Y$,
- $H_{/R}(x_{/R}) = \bigcup_{x \in x_{/R}} H(x)$.

Lemma 4.2.10 ([91])**.** *Let S be an autonomous metric system, R be an equivalence relation on X, and let the autonomous metric system $S_{/R}$ be the power quotient system of S by R. For any*

$$\varepsilon \geq \max_{\substack{x \in x_{/R} \\ x_{/R} \in X/R}} d(H(x), H_{/R}(x_{/R})),$$

with d the Hausdorff distance over the set 2^Y, $S_{/R}$ ε-approximately simulates S, i.e. $S \preceq^{\varepsilon}_{\mathscr{S}} S_{/R}$.

Now, we appropriately modify Definition 4.2.9 and Lemma 4.2.10 for the case that one can construct an over approximation of the power quotient system, namely $\bar{S}_{/R}$.

Definition 4.2.11. *(Approximate Power Quotient System [92]) Let $S = (X, X_0, U, \longrightarrow, Y, H)$ be a system, R be an equivalence relation on X, and $S_{/R} = (X_{/R}, X_{/R,0}, U_{/R}, \xrightarrow[/R]{}, Y_{/R}, H_{/R})$ be the power quotient of S by R. An approximate power quotient of S by R, denoted by $\bar{S}_{/R}$, is a system $(X_{/R}, X_{/R,0}, U_{/R}, \xrightarrow[/\bar{R}]{}, \bar{Y}_{/R}, \bar{H}_{/R})$ such that, $\xrightarrow[/\bar{R}]{} \supseteq \xrightarrow[/R]{}$, $\bar{Y}_{/R} \supseteq Y_{/R}$, and $\bar{H}_{/R}(x_{/R}) \supseteq H_{/R}(x_{/R})$, for all $x_{/R} \in X_{/R}$.*

Corollary 4.2.12 ([92])**.** *Let S be a metric system, R be an equivalence relation on X, and let the metric system $\bar{S}_{/R}$ be the approximate power quotient system of S by R. For any*

$$\varepsilon \geq \max_{\substack{x \in x_{/R} \\ x_{/R} \in X_{/R}}} d(H(x), \bar{H}_{/R}(x_{/R})),$$

with d the Hausdorff distance over the set 2^Y, $\bar{S}_{/R}$ ε-approximately simulates S, i.e. $S \preceq^{\varepsilon}_{\mathscr{S}} \bar{S}_{/R}$.

4.2.3. TIMED SAFETY AUTOMATON

In what follows, we present a formal definition for TSA. A TSA [95] is a directed graph extended with real-valued variables (called clocks) that model the logical clocks. We define C as a set of finitely many clocks. Clock constraints are used to restrict the behavior of the automaton. A clock constraint is a conjunctive formula of atomic constraints of the form $x \bowtie n$ or $x - y \bowtie n$ for $x, y \in C$, $\bowtie \in \{\leq, <, =, >, \geq\}$ and $n \in \mathbb{N}$. We use $\mathscr{B}(C)$ to denote the set of clock constraints.

Definition 4.2.13. *(Timed Safety Automaton [94]) A* timed safety automaton TSA *is a sextuple* $(L, \ell_0, \mathsf{Act}, C, E, \mathsf{Inv})$ *where*

- L *is a set of finitely many locations (or vertices),*
- $\ell_0 \in L$ *is the initial location,*
- Act *is the set of actions,*
- C *is a set of finitely many real-valued clocks,*
- $E \subseteq L \times \mathscr{B}(C) \times \mathsf{Act} \times 2^C \times L$ *is the set of edges,*
- $\mathsf{Inv} : L \to \mathscr{B}(C)$ *assigns invariants to locations.*

The location invariants are restricted to constraints of the form: $c \leq n$ or $c < n$, where c is a clock and n is a natural number.

4.2.4. PROBLEM STATEMENT

We now state the main problem considered in this chapter. Consider the system $S = (X, X_0, \varnothing, \longrightarrow, Y, H)$ where

- $X = \mathbb{R}^n$,
- $X_0 = \mathbb{R}^n$,
- $(x, x') \in \longrightarrow$ iff $\xi_x\big(\tau(x)\big) = x'$ given by (4.1)-(4.2), and (4.7),
- $Y \subset \mathbb{R}_{>0}$,
- $H : \mathbb{R}^n \to \mathbb{R}^+$ where $H(x) = \tau(x)$.

The output of the above system generates all possible sequences of inter-sample intervals of the concrete system (4.1)-(4.2) with the TM (4.7).

Problem 4.2.14. *Provide a construction of power quotient systems $S_{/\mathscr{P}}$ of systems S as defined above.*

Based on Definition 4.2.9, we propose to construct the system

$$S_{/\mathscr{P}} = (X_{/\mathscr{P}}, X_{/\mathscr{P},0}, \varnothing, \underset{/\mathscr{P}}{\longrightarrow}, Y_{/\mathscr{P}}, H_{/\mathscr{P}})$$

where

- $X_{/\mathscr{P}} = \mathbb{R}^n_{/\mathscr{P}} := \{\mathscr{R}_1, \ldots, \mathscr{R}_q\}$,
- $X_{/\mathscr{P},0} = \mathbb{R}^n_{/\mathscr{P}}$,
- $(x_{/\mathscr{P}}, x'_{/\mathscr{P}}) \in \underset{/\mathscr{P}}{\longrightarrow}$ if there exists $x \in x_{/\mathscr{P}}$, there exists $x' \in x'_{/\mathscr{P}}$ such that $\xi_x(H(x)) = x'$ as determined by (4.1)-(4.2),

- $Y_{/\mathscr{P}} \subset 2^Y \subset \mathbb{IR}^+$, where $\mathbb{IR}^+$ represents the set of closed intervals $[a,b]$ such that $0 < a \leq b$,
- $H_{/\mathscr{P}}(x_{/\mathscr{P}}) = \big[\min_{x \in x_{/\mathscr{P}}} H(x), \max_{x \in x_{/\mathscr{P}}} H(x)\big] := \big[\underline{\tau}_{x_{/\mathscr{P}}}, \bar{\tau}_{x_{/\mathscr{P}}}\big]$.

The equivalence relation $\mathscr{P}$ on $\mathbb{R}^n$ partitions the state space of S (i.e., the ETC system) into the set $X_{/\mathscr{P}}$ with a finite cardinality. However, since the exact construction of $S_{/\mathscr{P}}$ is in general impossible, we construct instead $\bar{S}_{/\mathscr{P}}$ (see Definition 4.2.11). Later on, it will be shown that the constructed $\bar{S}_{/\mathscr{P}}$ is equivalent to a TSA.

4.3. ABSTRACTION

In this subsection, we introduce the required steps in our framework to solve Problem 4.2.14 in the following order: (i) a suitable definition of the equivalence relation $\mathscr{P}$ on $\mathbb{R}^n$, (ii) a tractable approach to compute the output map $\bar{H}_{/\mathscr{P}}$ and its corresponding output set $\bar{Y}_{/\mathscr{P}}$, and (iii) a reachability-based analysis to derive the discrete transitions among abstract states $x_{/\mathscr{P}}$.

4.3.1. STATE SET

The approach to construct the state set mainly relies on an intuitive observation from the inequality (4.11).

Remark 4.3.1 (Conic construction of state set)**.** *Consider that the right-hand side of (4.11) is used to analyze the sampling behavior of the definition* (4.10)*. The sampling periods of all states, located on a line that passes through the origin excluding the origin itself, are lower bounded by the same quantity, i.e., $\tau'(x) = \tau'(\lambda x)$, for all $\lambda \neq 0$.*

It is not difficult to see that a proper approach to abstract the state space is via partitioning it into a finite number of convex polyhedral cones $\mathscr{R}_s$ (which are pointed at the origin), where $s \in \{1,\dots,q\}$, $\bigcup_{s=1}^{q} \mathscr{R}_s = \mathbb{R}^n$, and q is the cardinality of the state set. This state space abstraction technique is proposed by [102], dividing each of the angular spherical coordinates of $x \in \mathbb{R}^n$: $\theta_1,\dots,\theta_{n-2} \in [0,\pi]$, $\theta_{n-1} \in [-\pi,\pi]$ into $\bar{m}$ (not necessarily equidistant) intervals resulting in $q = \bar{m}^{(n-1)}$ conic regions. Furthermore, since the term $x^\top \Theta(t) x$ is quadratic in x, it is sufficient to only analyze half of the state space (e.g., by taking $\theta_{n-1} \in [0,\pi]$). Thus, the equivalence relation $\mathscr{P}$ to construct the abstraction is

$$(x,x') \in \mathscr{P} \Leftrightarrow \text{ there exists } s \in \{1,\dots,q\} \text{ such that } x, x' \in \mathscr{R}_s,$$

where q is the number of equivalence classes. In simple words, the equivalence classes of $\mathscr{P}$ are defined by polyhedral cones pointed at the origin given by $\mathscr{R}_s = \{x \in \mathbb{R}^2 : x^\top Q_s x \geq 0\}$, $Q_s \in \mathscr{S}_2$ whenever $n = 2$ or $\mathscr{R}_s = \{x \in \mathbb{R}^n : E_s x \geq 0\}$, $E_s \in \mathscr{S}_{n \times p}$ otherwise.

4.3.2. OUTPUT MAP

In this subsection, we present the approach to construct the output map $\bar{H}_{/\mathscr{P}}$ and the output set $\bar{Y}_{/\mathscr{P}}$. Recall that for all $x \in \mathscr{R}_s$, the output $y_{/\mathscr{P}} \in \bar{Y}_{/\mathscr{P}}$ where $y_{/\mathscr{P}} = \bar{H}_{/\mathscr{P}}(x)$ simply denotes the time interval $[\underline{\tau}_s, \bar{\tau}_s]$. This time interval in turn implies that the sampling period $\tau(x) \in [\underline{\tau}_s, \bar{\tau}_s]$. Our approach is inspired by the polytopic embedding technique

proposed in [103]. An intuitive description of this technique follows. Consider the space of real matrices. This technique construct a sequence of convex polytopes around the matrix $\Phi(t)$. Notice that each vertex is a matrix here. (We shall denote the set of these vertices $\Phi_{\kappa,s}$ in the following.) By doing so, one replaces the evaluation of (4.9), which has to be done at infinitely many instants t, with the evaluation of $\Phi_{\kappa,s}$, which has to be done at finitely many vertices $\Phi_{\kappa,s}$.

Assume the existence of a scalar $\sigma > 0$ that denotes a time instant for which the TM (4.7) is enabled in the whole state space, i.e., $\Phi(t) \succeq 0$. Consider $N_{\text{conv}}+1$ is the number of vertices employed to define the polytope containing $\Phi(t)$ in a given time interval, and $\ell \geq 1$ denotes the number of time subdivisions considered in the time interval $[0,\sigma]$.

Lemma 4.3.2 (Lower bound of inter-sample interval)**.** *Let $s \in \{1,\dots,q\}$. Consider a time instant $\underline{\tau}_s \in (0,\sigma]$, a scalar μ and a symmetric matrix Ψ satisfying* (4.8a).

If $\underline{\Phi}_{(i,j),s} \preceq 0$ holds for all $(i,j) \in \mathcal{K}_s = \big(\{0,\dots,N_{\text{conv}}\} \times \{0,\dots,\lfloor \frac{\underline{\tau}_s \ell}{\sigma} \rfloor\}\big)$, it follows that $\Phi(t) \preceq 0$ for all $t \in [0,\underline{\tau}_s]$, with Φ defined in (4.8c) *and*

$$\underline{\Phi}_{(i,j),s} = \underline{\tilde{\Phi}}_{(i,j),s} + \eta I$$

$$\underline{\tilde{\Phi}}_{(i,j),s} = \begin{cases} \Sigma_{k=0}^{i} \underline{\hat{\Phi}}_{(i,j),s} \big(\frac{\sigma}{\ell}\big)^k, & \text{for } j < \Big\lfloor \frac{\underline{\tau}_s \ell}{\sigma} \Big\rfloor \\ \Sigma_{k=0}^{i} \underline{\hat{\Phi}}_{(i,j),s} \big(\underline{\tau}_s - j\frac{\sigma}{\ell}\big)^k, & \text{for } j = \Big\lfloor \frac{\underline{\tau}_s \ell}{\sigma} \Big\rfloor, \end{cases} \tag{4.13a}$$

$$\underline{\hat{\Phi}}_{(0,j),s} = \begin{bmatrix} L_{0,j} & \check{\Pi}_j^\top M^\top \\ M\check{\Pi}_j & -\Psi \end{bmatrix}, \underline{\hat{\Phi}}_{(k\geq 1,j),s} = \begin{bmatrix} L_{k,j} & \hat{\Pi}_j^\top \frac{(A^{k-1})^\top}{k!} M^\top \\ M\frac{A^{k-1}}{k!}\hat{\Pi}_j & 0 \end{bmatrix}, \tag{4.13b}$$

and

$$L_{0,j} = \check{\Pi}_j^\top M \check{\Pi}_j - N + \tilde{L}_{0,j} \tag{4.13c}$$

with

$$\tilde{L}_{0,j} = \begin{cases} W\mu \frac{\lambda_{\max}\left(E^\top E\right)}{\lambda_{\max}^A} \big(j\frac{\sigma}{\ell}\big)\big(e^{\lambda_{\max}^A j\frac{\sigma}{\ell}} - 1\big) I, & \text{for} \quad \lambda_{\max}^A \neq 0, \\ W\mu \frac{\lambda_{\max}\left(E^\top E\right)}{\lambda_{\max}^A} \big(j\frac{\sigma}{\ell}\big)^2 I, & \text{for} \quad \lambda_{\max}^A = 0, \end{cases} \tag{4.13d}$$

$$L_{1,j} = \check{\Pi}_j^\top M \hat{\Pi}_j + \hat{\Pi}_j^\top M \check{\Pi}_j + \tilde{L}_{1,j} \tag{4.13e}$$

with

$$\tilde{L}_{1,j} = \begin{cases} W\mu \frac{\lambda_{\max}(E^\top E)}{\lambda_{\max}^A} \Big[\big(j\frac{\sigma}{\ell}\big) e^{\lambda_{\max}^A j\frac{\sigma}{\ell}} \lambda_{\max}^A + e^{\lambda_{\max}^A j\frac{\sigma}{\ell}} - 1\Big] I, & \text{for } \lambda_{\max}^A \neq 0, \\ W\mu(2j\frac{\sigma}{\ell})\lambda_{\max}(E^\top E) I, & \text{for } \lambda_{\max}^A = 0, \end{cases} \tag{4.13f}$$

$$L_{2,j} = \check{\Pi}_j^\top M \frac{A}{2!} \hat{\Pi}_j + \hat{\Pi}_j^\top \frac{A^\top}{2!} M \check{\Pi}_j + \hat{\Pi}_j^\top M \hat{\Pi}_j + \tilde{L}_{2,j} \tag{4.13g}$$

with

$$\tilde{L}_{2,j} = \begin{cases} W\mu \frac{\lambda_{\max}(E^\top E)}{\lambda_{\max}^A} \Big[\big(j\frac{\sigma}{\ell}\big) e^{\lambda_{\max}^A j\frac{\sigma}{\ell}} \frac{(\lambda_{\max}^A)^2}{2!} + e^{\lambda_{\max}^A j\frac{\sigma}{\ell}} \lambda_{\max}^A\Big] I, & \text{for } \lambda_{\max}^A \neq 0, \\ W\mu\lambda_{\max}(E^\top E) I, & \text{for } \lambda_{\max}^A = 0, \end{cases} \tag{4.13h}$$

$$L_{k\geq 3,j} = \check{\Pi}_j^\top M \frac{A^{k-1}}{k!}\hat{\Pi}_j + \hat{\Pi}_j^\top \frac{(A^{k-1})^\top}{k!} M\check{\Pi}_j + \hat{\Pi}_j^\top \Big(\Sigma_{i=1}^{k-1} \frac{(A^{i-1})^\top}{i!} M \frac{A^{k-i-1}}{(k-i)!}\Big)\hat{\Pi}_j + \tilde{L}_{k,j} \tag{4.13i}$$

with

$$\tilde{L}_{k\geq 3,j} = \begin{cases} W\mu \dfrac{\lambda_{\max}(E^\top E)}{\lambda^A_{\max}} \Big[\big(j\dfrac{\sigma}{\ell}\big) e^{\lambda^A_{\max} j\frac{\sigma}{\ell}} \dfrac{(\lambda^A_{\max})^k}{k!} + e^{\lambda^A_{\max} j\frac{\sigma}{\ell}} \dfrac{(\lambda^A_{\max})^{k-1}}{(k-1)!} \Big] I, & \text{for } \lambda^A_{\max} \neq 0, \\ 0, & \text{for } \lambda^A_{\max} = 0, \end{cases} \tag{4.13j}$$

$$\begin{aligned} &\check{\Pi}_j = \check{F}_j(A-BK), \quad \hat{\Pi}_j = \hat{F}_j(A-BK), \\ &\check{F}_j = \int_0^{j\frac{\sigma}{\ell}} e^{As} ds, \qquad\qquad \hat{F}_j = A\check{F}_j + I, \end{aligned} \tag{4.13k}$$

and

$$\eta \geq \max_{t'\in[0,\frac{\sigma}{\ell}], r\in\{0,\ldots,\ell-1\}} \lambda_{\max}\Big(\Phi(t'+r\frac{\sigma}{\ell}) - \Sigma_{k=0}^{N} \underline{\hat{\Phi}}_{k,r}(t')^k\Big). \tag{4.14}$$

Then, using the S-procedure, the following theorem provides an approach to regionally reduce the conservatism involved in the estimates $\underline{\tau}_s$ obtained from Lemma 4.3.2.

Theorem 4.3.3 (Regional lower bound of inter-sample interval)**.** *Consider a scalar $\underline{\tau}_s \in (0,\sigma]$, a scalar μ and a symmetric matrix Ψ satisfying (4.8a), and matrices $\underline{\Phi}_{\kappa,s}$, $\kappa=(i,j)\in\mathcal{K}_s$, defined as in Lemma 4.3.2.*

If there exist scalars $\underline{\alpha}_{\kappa,s}\geq 0$ (for $n=2$) or symmetric matrices $\underline{U}_{\kappa,s}$ with nonnegative entries (for $n\geq 3$) such that for all $\kappa\in\mathcal{K}_s$ the LMIs

$$\begin{cases} \underline{\Phi}_{(i,j),s} + \begin{bmatrix} \underline{\alpha}_{(i,j),s} Q_s & 0 \\ 0 & 0 \end{bmatrix} \preceq 0, & \text{for } n=2, \\ \underline{\Phi}_{(i,j),s} + \begin{bmatrix} E_s^T \underline{U}_{(i,j),s} E_s & 0 \\ 0 & 0 \end{bmatrix} \preceq 0, & \text{for } n\geq 3, \end{cases} \tag{4.15}$$

hold, the inter-sample interval (4.7) of the system (4.1)-(4.2) is then regionally bounded from below by $\underline{\tau}_s$, for all $x\in\mathscr{R}_s$.

One can follow a similar approach to find the upper bounds $\bar{\tau}_s$ on the inter-sample times that is outlined in Lemma 4.3.4 and Theorem 4.3.5.

Lemma 4.3.4 (Upper bound of inter-sample interval)**.** *Let $s\in\{1,\ldots,q\}$. Consider a time instant $\bar{\tau}_s\in[\underline{\tau}_s,\sigma]$, a scalar μ and a matrix Ψ satisfying the LMI conditions given in Lemma 4.3.2.*

If $\bar{\Phi}_{(i,j),s}\preceq 0$ holds for all $(i,j)\in\mathcal{K}_s=\big(\{0,\ldots,N_{\text{conv}}\}\times\{\lfloor\frac{\bar{\tau}_s\ell}{\sigma}\rfloor,\ldots,\ell-1\}\big)$, then, it follows that $\Phi(t)\succeq 0$ for all $t\in[\bar{\tau}_s,\sigma]$, with Φ defined in (4.8c) and

$$\bar{\Phi}_{(i,j),s} = -\bar{\tilde{\Phi}}_{(i,j),s} - \eta I,$$

$$\bar{\bar{\Phi}}_{(i,j),s} = \begin{cases} \sum_{k=0}^{i} L_{k,j}\Big(\dfrac{(j+1)\sigma}{\ell} - \bar{\tau}_s\Big)^k, & \text{for } j = \Big\lfloor \dfrac{\bar{\tau}_s \ell}{\sigma} \Big\rfloor, \\ \sum_{k=0}^{i} L_{k,j}\big(\dfrac{\sigma}{\ell}\big)^k, & \text{for } j > \Big\lfloor \dfrac{\bar{\tau}_s \ell}{\sigma} \Big\rfloor, \end{cases}$$

where $L_{k,j}$ are given by (4.13c)-(4.13k) *and η is defined in (4.14).*

Theorem 4.3.5 (Regional upper bound of inter-sample interval)**.** *Consider a scalar $\bar{\tau}_s \in [\underline{\tau}_s, \sigma]$, a scalar μ and a symmetric matrix Ψ satisfying (4.8a), and matrices $\bar{\Phi}_{\kappa,s}$, $\kappa = (i,j) \in \mathscr{K}_s$, defined as in Lemma 4.3.4.*

If there exist scalars $\bar{\alpha}_{\kappa,s} \geq 0$ (for $n = 2$) or symmetric matrices $\bar{U}_{\kappa,s}$ with nonnegative entries (for $n \geq 3$) such that for all $\kappa \in \mathscr{K}_s$ the LMIs

$$\begin{cases} \bar{\Phi}_{(i,j),s} - \begin{bmatrix} \bar{\alpha}_{(i,j),s} Q_s & 0 \\ 0 & 0 \end{bmatrix} \preceq 0, & \text{for } n = 2, \\ \bar{\Phi}_{(i,j),s} - \begin{bmatrix} E_s^T \bar{U}_{(i,j),s} E_s & 0 \\ 0 & 0 \end{bmatrix} \preceq 0, & \text{for } n \geq 3, \end{cases} \tag{4.16}$$

hold, the inter-sample interval (4.7) *of the system* (4.1)-(4.2) *is then regionally bounded from above by $\bar{\tau}_s$, for all $x \in \mathscr{R}_s$.*

4.3.3. TRANSITION RELATIONS

We next introduce the approach to find the transitions in $\bar{S}_{/\mathscr{P}}$. To this end, it is required to compute the reachable set of each cone $\mathscr{R}_s$ over the time interval $[\underline{\tau}_s, \bar{\tau}_s]$. In the sequel, we present how one is able to compute an over-approximation of the reachable set of each cone by the Minkowski sum of two sets.

Recall that the evolution of states over the time interval $[\underline{\tau}_s, \bar{\tau}_s]$ is given by $\xi_x(\tau) = \Lambda(\tau)x + \Omega(\tau)$. Let $X_{0,s}$ be the set that the initial states lie in. Denote by $\mathscr{X}_{[\underline{\tau}_s, \bar{\tau}_s]}(X_{0,s})$ the reachable set of $X_{0,s}$ during the time interval $[\underline{\tau}_s, \bar{\tau}_s]$, i.e.,

$$\big\{x' \in \mathbb{R}^n : \text{ there exists } x \in X_{0,s}, \text{ there exists } \tau \in [\underline{\tau}_s, \bar{\tau}_s], x' = \xi_x(\tau)\big\}.$$

Furthermore, define

$$\mathscr{X}^1_{[\underline{\tau}_s, \bar{\tau}_s]}(X_{0,s}) := \big\{x' \in \mathbb{R}^n : \text{ there exists } x \in X_{0,s}, \text{ there exists } \tau \in [\underline{\tau}_s, \bar{\tau}_s], x' = \Lambda(\tau)x\big\},$$
$$\mathscr{X}^2_{[\underline{\tau}_s, \bar{\tau}_s]}(X_{0,s}) := \big\{x' \in \mathbb{R}^n : \text{ there exists } x \in X_{0,s}, \text{ there exists } \tau \in [\underline{\tau}_s, \bar{\tau}_s], x' = \Omega(\tau)\big\}.$$

Thus,

$$\mathscr{X}_{[\underline{\tau}_s, \bar{\tau}_s]}(X_{0,s}) := \mathscr{X}^1_{[\underline{\tau}_s, \bar{\tau}_s]}(X_{0,s}) \bigoplus \mathscr{X}^2_{[\underline{\tau}_s, \bar{\tau}_s]}(X_{0,s})$$

A suitable approach to define the initial set $X_{0,s}$ is now introduced. In [91, Section III.B.3], it has been shown that it is enough to consider subsets $X_{0,s} \subset \mathscr{R}_s$ being convex polytopes with each vertex placed on each of the extreme rays of $\mathscr{R}_s$ (excluding the origin) to compute $\mathscr{X}^1_{[\underline{\tau}_s, \bar{\tau}_s]}(X_{0,s})$. Then, one can effectively compute an over-approximation

of the reachable set of a polytope under LTI dynamics, denoted by $\hat{\mathscr{X}}^1_{[\underline{\tau}_s,\bar{\tau}_s]}(X_{0,s})$, see e.g., [104]. Furthermore, one has

$$\begin{aligned}\|\Omega(\tau)\| &= \|\int_0^\tau e^{A(\tau-s)}E\omega(s)ds\| \\ &\leq \int_0^\tau \|e^{A(\tau-s)}E\omega(s)\|ds \\ &\leq \int_0^\tau \|e^{A(\tau-s)}\|\|E\||\omega(s)|ds \\ &\leq W|x|\|E\|\int_0^\tau |e^{\mu(A)(\tau-s)}|ds \\ &= \rho(\tau)|x|,\end{aligned}$$

where $\rho(\tau) = W\|E\|\int_0^\tau |e^{\mu(A)(\tau-s)}|ds$. As a result, $\mathscr{X}^2_{[\underline{\tau}_s,\bar{\tau}_s]}(X_{0,s})$ can be over-approximated by a second order cone

$$\hat{\mathscr{X}}^2_{[\underline{\tau}_s,\bar{\tau}_s]}(X_{0,s}) := \{x' \in \mathbb{R}^n : \text{ there exists } x \in X_{0,s}, \text{ there exists } \tau \in [\underline{\tau}_s,\bar{\tau}_s], |x'| \leq \rho(\bar{\tau}_s)|x|\}.$$

To compute the transitions in $\bar{S}_{/\mathscr{P}}$, it thus suffices to derive the intersection between the over-approximation

$$\hat{\mathscr{X}}_{[\underline{\tau}_s,\bar{\tau}_s]}(X_{0,s}) := \hat{\mathscr{X}}^1_{[\underline{\tau}_s,\bar{\tau}_s]}(X_{0,s}) \bigoplus \hat{\mathscr{X}}^2_{[\underline{\tau}_s,\bar{\tau}_s]}(X_{0,s})$$

and all the conic regions $\mathscr{R}_t$ where $t \in \{1,\ldots,q\}$. To the compute transitions, it is required to check whether the convex feasibility problem

$$\mathscr{R}_t \cap \hat{\mathscr{X}}_{[\underline{\tau}_s,\bar{\tau}_s]}(X_{0,s}) \neq \varnothing, \tag{4.17}$$

for each conic region $\mathscr{R}_t$ hold, which can be solved by existing convex analysis tools. There exists a transition from abstract state $\mathscr{R}_s$ to $\mathscr{R}_t$ in $\bar{S}_{/\mathscr{P}}$ in the case that (4.17) is satisfied.

4.3.4. TIMED SAFETY AUTOMATA REPRESENTATION

In this subsection, we first point out the connection between an abstract state $x_{/\mathscr{P}} \in X_{/\mathscr{P}}$ and its corresponding output $y_{/\mathscr{P}} \in \bar{Y}_{/\mathscr{P}}$ [91]. The system $\bar{S}_{/\mathscr{P}}$

1. remains at $x_{/\mathscr{P}}$ during the time interval $[0, \underline{\tau}_{x_{/\mathscr{P}}})$,
2. possibly leaves $x_{/\mathscr{P}}$ during the time interval $[\underline{\tau}_{x_{/\mathscr{P}}}, \bar{\tau}_{x_{/\mathscr{P}}})$, and
3. is forced to leave $x_{/\mathscr{P}}$ at the time instant $\bar{\tau}_{x_{/\mathscr{P}}}$.

Thus, the semantics of $\bar{S}_{/\mathscr{P}}$ is equivalent to a timed safety automaton

$$\mathsf{TSA} = (L, \ell_0, \mathsf{Act}, C, E, \mathsf{Inv}),$$

where

- $L = X_{/\tilde{\mathscr{P}}}$,
- $\ell_0 := \mathscr{R}_s$ such that $\xi(0) \in \mathscr{R}_s$,
- Act $= \{*\}$ is an arbitrary symbol,
- $C = \{c\}$,
- E is given by all tuples $(\mathscr{R}_s, g, a, r, \mathscr{R}_t)$ such that $(\mathscr{R}_s, \mathscr{R}_t) \in \xrightarrow[/\tilde{\mathscr{P}}]{}$, $g = \{c : c \in [\underline{\tau}_s, \bar{\tau}_s]\}$, $a = *$, and r is given by $c := 0$,
- $\mathsf{Inv}(\mathscr{R}_s) := \{c : c \in [0, \bar{\tau}_s]\}$, for all $s \in \{1, \ldots, q\}$.

Remark 4.3.6 (Drawback of construction technique)**.** *Although the construction technique presented in this section is offline, it is exponentially dependent on $n-1$ (where n is the number of states). Hence, the proposed construction technique becomes computationally expensive for higher-order systems.*

4

4.4. TECHNICAL PROOFS

4.4.1. PROOF OF LEMMA 4.3.2

We first divide the time interval $[0, \sigma]$ into ℓ subintervals. The reason behind this subdivision is to reduce the conservatism in polytopic embedding. Let the instant $t \in [0, \sigma]$. Define $t = t' + j\frac{\sigma}{\ell}$ where $t' \in [0, \chi]$, with $\chi = \frac{\sigma}{\ell}$ for $j < \lfloor \frac{\underline{\tau}_s \ell}{\sigma} \rfloor$ and $\chi = \underline{\tau}_s - j\frac{\sigma}{\ell}$ otherwise. Denote $\Lambda(t) - I$ by $\mathscr{X}(t)$. One has

$$\mathscr{X}(t) = \left[\int_0^{j\frac{\sigma}{\ell}} e^{As} ds + \int_0^{t'} e^{As} ds \Big(A \int_0^{j\frac{\sigma}{\ell}} e^{As} ds + I \Big) \right] (A - BK).$$

Rewrite the above equality in a more compact form

$$\mathscr{X}(t) = \check{\Pi}_j + \int_0^{t'} e^{As} ds \hat{\Pi}_j.$$

Substitute the compact form of $\mathscr{X}(t)$ into (4.8c) and as a result,

$$\begin{aligned}
\Phi_{11}(t) &= \check{\Pi}_j^\top M \check{\Pi}_j + \check{\Pi}_j^\top M \Big(\int_0^{t'} e^{As} ds \Big) \hat{\Pi}_j + \hat{\Pi}_j^\top \Big(\int_0^{t'} e^{As} ds \Big)^\top M \check{\Pi}_j \\
&\quad + \hat{\Pi}_j^\top \Big(\int_0^{t'} e^{As} ds \Big)^\top M \Big(\int_0^{t'} e^{As} ds \Big) \hat{\Pi}_j + t W \mu \lambda_{\max}(E^\top E) d_A(t) I - N, \\
\Phi_{12}(t) &= \check{\Pi}_j^\top M + \hat{\Pi}_j^\top \Big(\int_0^{t'} e^{As} ds \Big)^\top M, \\
\Phi_{21}(t) &= \Phi_{12}^\top(t), \\
\Phi_{22}(t) &= -\Psi.
\end{aligned}$$

We next use the polytopic embedding approach proposed by [103] to abstract away t in (4.8c). In the polytopic embedding approach, the underlying idea is as follows. We

first approximate the matrix functionals $tW\mu\lambda_{\max}\big(E^\top E\big)d_A(t)I$ and $\int_0^{t'} e^{As}ds$ in $\Phi(t)$ by their N_{conv}-th order Taylor series expansions.

In the case of $tW\mu\lambda_{\max}\big(E^\top E\big)d_A(t)I$, notice that the matrices $L_{i,j}$ represents the Taylor expansion's terms, where $(i,j)\in\big(\{0,\dots,N_{\text{conv}}\}\times\{\lfloor\frac{\bar{\tau}_s\ell}{\sigma}\rfloor,\dots,\ell-1\}\big)$. Also, notice that

$$\int_0^{t'} e^{As}ds \simeq \Sigma_{i=1}^{N_{\text{conv}}}\frac{A^{i-1}}{i!}(t')^i.$$

As a result of these approximations, we take into account the introduced error. Let η be an upper bound on this error. The procedure to find η follows. The exact Taylor expansion of $\Phi(t)$ is given by $\Sigma_{k=0}^{\infty}\hat{\underline{\Phi}}_{k,j}(t')^k$ where $\hat{\underline{\Phi}}_{k,j}$ is given in (4.13b). However, we consider the N_{conv}-th order expansion of $\Phi(t)$ and denote it by $\tilde{\underline{\Phi}}_{(N_{\text{conv}},j)}(t')$, i.e., $\tilde{\underline{\Phi}}_{(N_{\text{conv}},j)}(t')=\Sigma_{k=0}^{N_{\text{conv}}}\hat{\underline{\Phi}}_{k,j}(t')^k$. The error introduced by the approximation is

$$\underline{R}_{(N_{\text{conv}},j)}(t')=\Phi(t)-\tilde{\underline{\Phi}}_{(N_{\text{conv}},j)}(t')$$

which is a symmetric matrix. One is able to derive an upper bound on $\underline{R}_{(N_{\text{conv}},j)}(t')\preceq\eta I$ where the scalar η is given by (4.14). Observe that $\tilde{\underline{\Phi}}_{(N_{\text{conv}},j)}(t')+\eta I\preceq 0$ implies $\Phi(t)\preceq 0$. The function $\tilde{\underline{\Phi}}_{(N_{\text{conv}},j)}(\cdot)+\eta I$ is a polynomial function in its argument.

Next, we use the convex embedding technique in [103] to show that $\underline{\Phi}_{(i,j),s}\preceq 0$, for all $(i,j)\in\mathscr{K}_s$, with $\underline{\Phi}_{(i,j),s}=\Sigma_{k=0}^{i}L_{k,j}\chi^k+\eta I$ implies $(\tilde{\underline{\Phi}}_{N_{\text{conv}},j}(\sigma')+\eta I)\preceq 0$. As discussed above, $(\tilde{\underline{\Phi}}_{N_{\text{conv}},j}(\sigma')+\eta I)\preceq 0$ leads to $\Phi(t)\preceq 0$, for all $t\in[0,\underline{\tau}_s]$. This concludes the proof.

4.4.2. PROOF OF THEOREM 4.3.3

Consider scalars $\underline{\alpha}_{(i,j),s}$ for $n=2$ (or matrices $\underline{U}_{(i,j),s}$ for $n\geq 3$) satisfying LMI conditions given in (4.15) for $s\in\{1,\dots,q\}$. By the virtue of Schur complement and Lemma 4.3.2, it follows that $\Phi(t)+\underline{\alpha}_{(i,j),s}Q_s\preceq 0$ for $n=2$ (or $\Phi(t)+E_s^\top\underline{U}_{(i,j),s}E_s\preceq 0$ for $n\geq 3$). Then, since for all $x\in\mathscr{R}_s$, $\{x\in\mathbb{R}^2:\ x^\top Q_s x\geq 0\}$ for $n=2$ (or $\{x\in\mathbb{R}^n:\ E_s x\geq 0\}$ for $n\geq 3$), the S-procedure implies that $x^\top\Phi(t)x\leq 0$, for all $t\in[0,\underline{\tau}_s]$. Finally, Theorem 4.2.4 guarantees that for all $x\in\mathscr{R}_s$, the inter-sample time $\tau(x)$ is lower bounded by $\underline{\tau}_s$.

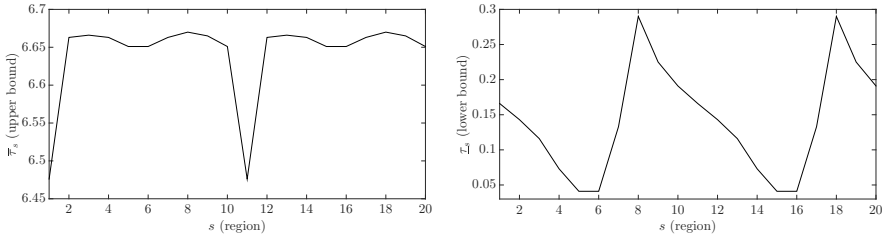
4.4.3. SKETCHES OF PROOFS OF LEMMA 4.3.4 & THEOREM 4.3.5

A sketch of proof is given. The polytopic embedding according to time of $-\Phi(t)$ enable us to show that $-\Phi(t)\preceq 0$ (or $\Phi(t)\succeq 0$) if $\bar{\Phi}_{\kappa,s}\preceq 0$, for all $\kappa\in\mathscr{K}_s$. Then, by applying the Schur complement on $-\Phi(t)$, it follows $-\Theta(t)\preceq 0$ (or $\Theta(t)\succeq 0$) and as a result $-x^\top\Theta(t)x\leq 0$. Furthermore, considering (4.11) in Theorem 4.2.4, i.e., $-\mathscr{F}_\omega(x,t)\geq -x^\top\Theta(t)x$, the claims in Lemma 4.3.4 and Theorem 4.3.5 follow.

4.5. NUMERICAL EXAMPLE

We illustrate the effectiveness of the theoretical results of this chapter in a numerical example. Consider an LTI system, used as an example in [47], and add a perturbation term $\omega(t)$,

$$\dot{\xi}(t)=\begin{bmatrix}0&1\\-2&3\end{bmatrix}\xi(t)+\begin{bmatrix}0\\1\end{bmatrix}\nu(t)+\begin{bmatrix}0\\1\end{bmatrix}\omega(t), \tag{4.18}$$



(a) Upper bounds.

(b) Lower bounds.

Figure 4.1: (Left) Upper bounds on regional inter-sample times. (Right) Lower bounds on regional inter-sample times.

where the perturbation bound $W = 0.001$. We set the scalars associated with $\mathscr{L}_2$-based TM as follows: $\gamma = 100$ and $\beta = 0.25$, see (4.6) and (4.7). Then, solving the ARE associated with $\mathscr{L}_2$ stability, the control update law (implemented in a sample-and-hold fashion) is computed, that is

$$v(t) = -K\xi(t_k) = -[0.2361 \quad 6.2367]\xi(t_k), \text{ for all } t \in [t_k, t_{k+1}),$$

where t_k denotes the sampling instants and $k \in \mathbb{N} \cup \{0\}$. We set the order of polynomial approximation $N_{\text{conv}} = 7$, the number of polytopic subdivisions $\ell = 800$, the upper bound of the inter-sample intervals $\sigma = 8$, and the number of angular sub-divisions $\bar{m} = 10$, thus, $q = 2 \times 10^{(2-1)} = 20$. Applying the results from Section 4.3.2, we get the precision of abstraction $\varepsilon = 6.100$. This precision is relatively large with respect to the one derived in [92], that is $\varepsilon = 0.119$. The reason for such a large value for the perturbed case is as follows. The possible stabilizing effect of disturbance on the dynamics (4.1) can in fact enhance the stability of control system between two, consecutive triggering instants. As a result, the derived $\bar{\tau}_s$ possesses a relatively larger value compared to the one in the unperturbed case. In Figure 4.1, the derived lower and upper bounds are depicted. It is evident that the regional lower bounds $\underline{\tau}_s$ are less conservative compared to the minimal inter-sample time. One can effectively employ these lower bounds as less conservative measures for scheduling purposes. Figure 4.2 represents the conic regions s and the associated $\underline{\tau}_s$ and $\bar{\tau}_s$. In addition, Figure 4.3 depicts the simulation of the control system for a simulation time of 15 sec. It is clear that the bounds derived by the analysis given in Section 4.3.2 have been respected by the sampling periods generated by the control system. Figure 4.4 finally depicts the result of applying the procedure introduced in Section 4.3.3 to derive all the transitions in the TSA.

4.6. CONCLUSIONS

Most of the existing ETC strategies are equipped with a quantity, the so-called minimum inter-sample time, that indicates the maximum utilization of communication bandwidth. Despite the fact that such a quantity can practically be used in scheduling of ETC feedback loops, it does not allow us to exploit the beneficiary features of ETC strategies in the context of scheduling. In this chapter, we presented an approach to capture the sampling behavior of perturbed LTI systems with $\mathscr{L}_2$-based triggering mechanisms by

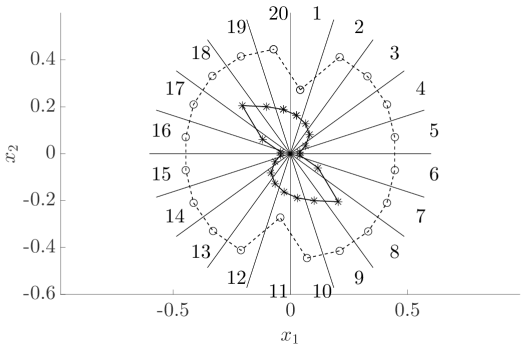


Figure 4.2: Polar representation of bounds. The radial distance from the origin of each asterisk indicates the regional lower bound of the indexed cone. Furthermore, in the case of circles, the distance indicates the regional upper bound of the indexed cone minus 6.2 sec, i.e., $\bar{\tau}_s - 6.2$ sec (for the sake of clarity of the figure).

timed safety automata. We formally established that the derived timed automaton ε-approximately simulates the ETC system under investigation. As a result of such a derivation, an object is constructed that can be used to synthesize scheduling policies for ETC feedback loops (i.e., the main contribution of this chapter).

In what follows, we list several future research directions to extend the results of this chapter. Exploiting already available tools for the synthesis of timed automata, one can employ the above results to synthesize conflict-free policies in WNCSs, see e.g., [91] which proposed a centralized scheduling of feedback policies. Another promising direction to follow is to find a fully decentralized approach instead of the centralized approach proposed in [91]. The next research direction concerns decentralized ETC strategies in the literature. In these studies, the existence of a minimum inter-sample time among different subsystems is commonly absent (to the best of author's knowledge), see e.g., [105]. The lack of such a quantity forces scheduler designers to add extra provisions in the scheduling policy. It is thus interesting to investigate the possibility of deriving cross-subsystem inter-sample times by modifying the framework used in this chapter. A successful answer to this research hypothesis basically guarantees conflict-free triggering instants generated by multiple subsystems. We finally remark a modification that is related to the regional upper bounds derived in Section 4.3.4. In the numerical example, it has been observed that the upper bounds are relatively large because of possible stabilizing effects of perturbations on the error dynamics. These large upper bounds may result in a timed automaton with a huge number of transitions. As a consequence, the schedulability of such an abstraction becomes a demanding task. To circumvent this issue, one can arbitrate a user-defined upper bound on the triggering mechanism to facilitate the scheduling process. For example in the case of WNCSs, this type of assumption is closely related to periodic-ETC strategies, see e.g., [106]. Moreover, the assumed bound can be seen as a sort of network heartbeat that forces ETC feedback loops to be updated.

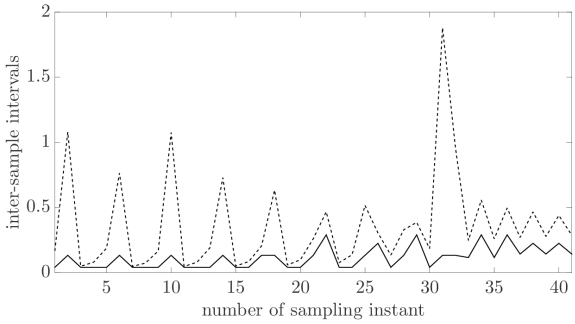


Figure 4.3: Validation of the derived lower bounds during the simulation. The solid line (dashed line) represents the derived lower bounds on inter-sample intervals (the generated inter-sample intervals by the ETC system during the simulation).

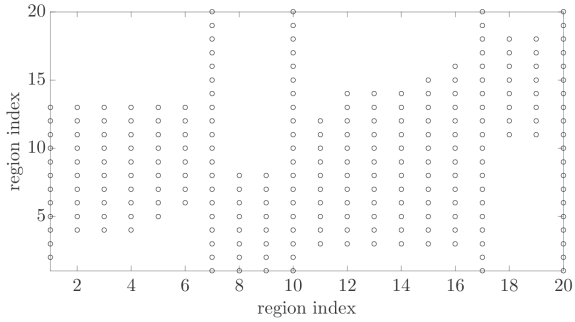


Figure 4.4: Schematic representation of all the edges in the timed automaton generated by the ETC system. Each circle at the coordinate (i, j) denotes an edge from location i to location j.

5

CONCLUSIONS

In this dissertation, three problems have been addressed that lie within the contexts of optimization and control. In the sequel, we summarize the proposed approaches for each problem instance followed by possible future research directions. We then provide two research directions that are motivated by the cross-chapter relations.

First problem:
Summary: Inspired by the recent viewpoint to optimization algorithms, we regarded an optimization algorithm as a controlled dynamical system. In this context, we set the desired performance measure to be an exponential rate of convergence. Two hybrid control frameworks were proposed to achieve the desired performance. The core idea behind the synthesis is based on a trajectory-based analysis under the assumption that the objective function satisfies the Polyak–Łojasiewicz inequality. In order to construct iterative algorithms, we next employed the forward-Euler method as the temporal discretization method. The conditions on the discretization step-size to guarantee an exponential convergence are further derived.

Future research directions: There are multiple directions in which the proposed approaches can be extended or modified.

- **(Modification of discretization)** As it has been pointed out in Chapter 2, the proposed discretization is naive, in the sense that it does not exploit the structure of the proposed input signals. We thus expect that the employment of a more advanced discretization method is a logical avenue to explore.

- **(Extension to other settings)** It is interesting to investigate the possibility of applying a similar type of analysis for other optimization settings, such as constrained or distributed cases.

Second problem:
Summary: The second problem considers the possibility of event-based implementation of an RMPC method. The principal idea of the proposed methodology is to exploit the optimal state and input trajectories computed at the last triggering instant. With this

mindset, we proposed two convex optimization formulations to identify a sequence of hyper-rectangles around the optimal state trajectories. This sequence of sets is then defined as the triggering mechanism in the following sense: as long as deviations of actual states, observed at sensory units and caused by model mismatches, are confined to this sequence, robust feasibility and stability of the closed-loop system are guaranteed. A distinctive feature of the proposed approach to design the triggering mechanism is its complete separation from the control synthesis part.

Future research directions: There are multiple directions that can be explored to extend the results of the proposed approach.

- **(Handling directional sensitivity)** One interesting extension is addressing the directional sensitivity of the proposed approach. At the moment, we assess two possibilities to address this issue. First, one can study the impacts of the nilpotent gains on the directional sensitivity with the hope of reducing the level of asymmetrical behavior. Motivated by threshold approaches in the literature, another possibility is to employ directionally-aware thresholds in the design process of the triggering mechanism.

- **(Extension to other settings)** One of the most interesting questions that comes to one's mind is to approach other types of RMPC methods with a similar viewpoint. Other types of straightforward extensions, such as considering a distributed setting, are also of interest.

Third problem:

Summary: To address the over-conservatism involved in the traditional (periodic) implementation of control actions, event-triggering control approaches provide a theoretically sound alternative. However, this alternative generally leads to a complicated sampling behavior that is difficult to be realized by real-time system engineers. In the third problem, we sought to construct a formal framework to translate the triggering behavior of perturbed LTI systems with an event-triggering implementation. To do so, an object, i.e., a timed safety automaton, is constructed which can be used to efficiently schedule event-triggering control tasks in a networked system. We used tools from stability analysis of delayed systems and reachability analysis of linear systems to construct the timed safety automaton.

Future research directions: There are multiple paths that can be followed to extend the derived results.

- **(Simplification of automaton)** Due to the perturbed nature of the model, it is highly probable that certain discrete states in the derived automaton possess a large number of out-going transitions. (This large number of out-going transitions stems from the fact that the upper-bounds of inter-execution times for the considered discrete states are relatively large.) Generally speaking, the occurrence of such a case is undesirable since the required scheduler becomes complicated. One possibility to circumvent this issue is to force a predefined upper-bound on the inter-execution times of all the discrete states.

- **(Extension to other settings)** A next natural step is the extension of the proposed approach to more general settings, such as nonlinear systems (that is already un-

der investigation by other researchers). In particular, the nonlinear systems, that have tractable reachability analysis, are the most suitable candidates for such an extension.

Inspired by cross-chapter relations, we next mention several research directions to extend the results of this dissertation.

- **(Dynamical systems view for optimization-based controllers)** As we have already discussed in Chapter 2, optimization algorithms are dynamical systems. As a result, one can view an online optimization-based control strategy as a dynamic controller. This controller is inter-connected to the plant via the feedback loop. In order to tackle complications that arise in such controllers, such a mindset provides us with tools from *singular perturbation theory* and *passivity theory* to hopefully propose a more unified approach.

- **(Timing abstraction of event-triggering MPC methods)** Following a similar approach proposed in Chapter 4, it is interesting to provide an object that can capture the timing behavior of event-triggering MPC methods. The main issue that can prevent us in doing so is the unavailability of the control law in a closed form. As a result, the reachability analysis becomes a rather difficult (if not impossible) task to do. On the other hand, set-based optimization methods are prime candidates to be equipped with this kind of timing behavior.

REFERENCES

[1] A. S. Kolarijani, P. Mohajerin Esfahani, and T. Keviczky, "Fast gradient-based methods with exponential rate: A hybrid control framework," in *Proceedings of the 35th International Conference on Machine Learning (ICML 2018)*, 2018.

[2] ——, "A hybrid control framework for fast methods under invexity: Non-Zeno trajectories with exponential rate," in *57th IEEE Conference on Decision and Control (CDC'18)*, 2018.

[3] ——, "Continuous-time accelerated methods via a hybrid control lens," *Submitted to IEEE Transactions on Automatic Control*, 2018.

[4] S. Bregman, A. S. Kolarijani, and T. Keviczky, "Robust model predictive control with aperiodic actuation," in *56th IEEE Conference on Decision and Control (CDC'17)*, 2017, pp. 5457–5462.

[5] A. S. Kolarijani, S. C. Bregman, P. Mohajerin Esfahani, and T. Keviczky, "Decentralized event-based policy to implement a robust model predictive control approach," *to be submitted to IEEE Transactions on Automatic Control*, 2018.

[6] A. S. Kolarijani, M. Mazo Jr., and T. Keviczky, "Timing abstraction of perturbed LTI systems with $\mathscr{L}_2$-based event-triggering mechanism," in *55th IEEE Conference on Decision and Control (CDC'16)*, 2016, pp. 1364–1369.

[7] D. Lashkari and P. Golland, "Convex clustering with exemplar-based models," in *Advances in Neural Information Processing Systems (NIPS 2008)*, 2008, pp. 825–832.

[8] L. Bottou, "Stochastic gradient learning in neural networks," *Proceedings of Neuro-Nımes*, vol. 91, no. 8, 1991.

[9] R. Salakhutdinov, S. T. Roweis, and Z. Ghahramani, "Optimization with EM and expectation-conjugate-gradient," in *Proceedings of the 20th International Conference on Machine Learning (ICML 2003)*, 2003, pp. 672–679.

[10] Z. Allen-Zhu, "Katyusha: The first direct acceleration of stochastic gradient methods," *arXiv preprint arXiv:1603.05953*, 2016.

[11] S. Becker, J. Bobin, and E. J. Candès, "Nesta: A fast and accurate first-order method for sparse recovery," *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 1–39, 2011.

[12] E. Ghadimi, I. Shames, and M. Johansson, "Multi-step gradient methods for networked optimization," *IEEE Transactions on Signal Processing*, vol. 61, no. 21, pp. 5417–5429, 2013.

[13] A. Cabot, "The steepest descent dynamical system with control. applications to constrained minimization," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 10, no. 2, pp. 243–258, 2004.

[14] B. T. Polyak, "Some methods of speeding up the convergence of iteration methods," *USSR Computational Mathematics and Mathematical Physics*, vol. 4, no. 5, pp. 1–17, 1964.

[15] Y. Nesterov, "A method of solving a convex programming problem with convergence rate $\mathcal{O}(1/k^2)$," in *Soviet Mathematics Doklady*, vol. 27, no. 2, 1983, pp. 372–376.

[16] ——, *Introductory lectures on convex optimization: a basic course*. Springer Science and Business Media, 2004.

[17] ——, "Smooth minimization of non-smooth functions," *Mathematical Programming*, vol. 103, no. 1, pp. 127–152, 2005.

[18] ——, "Gradient methods for minimizing composite functions," *Mathematical Programming*, vol. 140, no. 1, pp. 125–161, 2013.

[19] D. Drusvyatskiy, M. Fazel, and S. Roy, "An optimal first order method based on optimal quadratic averaging," *arXiv preprint arXiv:1604.06543*, 2016.

[20] S. Bubeck, Y. T. Lee, and M. Singh, "A geometric alternative to Nesterov's accelerated gradient descent," *arXiv preprint arXiv:1506.08187*, 2015.

[21] Y. Drori and M. Teboulle, "Performance of first-order methods for smooth convex minimization: a novel approach," *Mathematical Programming*, vol. 145, no. 1-2, pp. 451–482, 2014.

[22] L. Lessard, B. Recht, and A. Packard, "Analysis and design of optimization algorithms via integral quadratic constraints," *SIAM Journal on Optimization*, vol. 26, no. 1, pp. 57–95, 2016.

[23] W. Su, S. Boyd, and E. Candès, "A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights," in *Advances in Neural Information Processing Systems (NIPS 2014)*, 2014, pp. 2510–2518.

[24] ——, "A differential equation for modeling Nesterov's accelerated gradient method: theory and insights," *Journal of Machine Learning Research*, vol. 17, no. 153, pp. 1–43, 2016.

[25] A. Nemirovskii, D. B. Yudin, and E. R. Dawson, "Problem complexity and method efficiency in optimization," 1983.

[26] A. Wibisono, A. C. Wilson, and M. I. Jordan, "A variational perspective on accelerated methods in optimization," *Proceedings of the National Academy of Sciences*, vol. 113, no. 47, pp. E7351–E7358, 2016.

[27] A. C. Wilson, B. Recht, and M. I. Jordan, "A Lyapunov analysis of momentum methods in optimization," *arXiv preprint arXiv:1611.02635*, 2016.

[28] A. Megretski and A. Rantzer, "System analysis via integral quadratic constraints," *IEEE Transactions on Autom.*, vol. 42, no. 6, pp. 819–830, 1997.

[29] M. Fazlyab, A. Ribeiro, M. Morari, and V. M. Preciado, "Analysis of optimization algorithms via integral quadratic constraints: Non-strongly convex problems," *arXiv preprint arXiv:1705.03615*, 2017.

[30] B. Hu and L. Lessard, "Dissipativity theory for Nesterov's accelerated method," *arXiv preprint arXiv:1706.04381*, 2017.

[31] J. C. Willems, "Dissipative dynamical systems part i: General theory," *Archive for Rational Mechanics and Analysis*, vol. 45, no. 5, pp. 321–351, 1972.

[32] B. O'Donoghue and E. Candès, "Adaptive restart for accelerated gradient schemes," *Foundations of Computational Mathematics*, vol. 15, no. 3, pp. 715–732, 2015.

[33] A. Nemirovski, "Efficient methods in convex programming," 2005.

[34] M. Gu, L.-H. Lim, and C. J. Wu, "Parnes: a rapidly convergent algorithm for accurate recovery of sparse and approximately sparse signals," *Numerical Algorithms*, vol. 64, no. 2, pp. 321–347, 2013.

[35] G. Lan and R. Monteiro, "Iteration-complexity of first-order penalty methods for convex programming," *Mathematical Programming*, vol. 138, no. 1-2, pp. 115–139, 2013.

[36] R. Goebel, R. G. Sanfelice, and A. R. Teel, *Hybrid dynamical systems: modeling, stability, and robustness*. Princeton University Press, 2012.

[37] J.-P. Aubin, J. Lygeros, M. Quincampoix, S. Sastry, and N. Seube, "Impulse differential inclusions: a viability approach to hybrid systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 1, pp. 2–20, 2002.

[38] R. Goebel and A. R. Teel, "Solutions to hybrid inclusions via set and graphical convergence with stability theory applications," *Automatica*, vol. 42, no. 4, pp. 573–587, 2006.

[39] J. Lygeros, K. H. Johansson, S. N. Simic, J. Zhang, and S. Sastry, "Dynamical properties of hybrid automata," *IEEE Transactions on Automatic Control*, vol. 48, no. 1, pp. 2–17, 2003.

[40] M. A. Hanson, "On sufficiency of the Kuhn-Tucker conditions," *Journal of Mathematical Analysis and Applications*, vol. 80, no. 2, pp. 545–550, 1981.

[41] B. D. Craven and B. M. Glover, "Invex functions and duality," *Journal of the Australian Mathematical Society*, vol. 39, no. 1, pp. 1–20, 1985.

[42] H. Karimi, J. Nutini, and M. Schmidt, *Linear convergence of gradient and proximal-gradient methods under the Polyak-Łojasiewicz condition*. Springer International Publishing, 2016, pp. 795–811.

[43] Y. Nesterov and B. T. Polyak, "Cubic regularization of newton method and its global performance," *Mathematical Programming*, vol. 108, no. 1, pp. 177–205, 2006.

[44] H. S. Khalil, *Nonlinear systems*, 3rd ed. Prentice Hall, 2002.

[45] J. Baillieul and P. J. Antsaklis, "Control and communication challenges in networked real-time systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 9–28, 2007.

[46] W. P. M. H. Heemels, K. H. Johansson, and P. Tabuada, "An introduction to event-triggered and self-triggered control," in *51st IEEE Conference on Decision and Control (CDC'12)*, 2012, pp. 3270–3285.

[47] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1680–1685, 2007.

[48] A. Girard, "Dynamic triggering mechanisms for event-triggered control," *IEEE Transactions on Automatic Control*, vol. 60, no. 7, pp. 1992–1997, 2015.

[49] A. Anta and P. Tabuada, "To sample or not to sample: self-triggered control for nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2030–2042, 2010.

[50] C. Nowzari and J. Cortés, "Self-triggered coordination of robotic networks for optimal deployment," *Automatica*, vol. 48, no. 6, pp. 1077–1087, 2012.

[51] E. F. Camacho and C. B. Alba, *Model predictive control*. Springer Science & Business Media, 2013.

[52] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.

[53] D. Q. Mayne, "Model predictive control: recent developments and future promise," *Automatica*, vol. 50, no. 12, pp. 2967–2986, 2014.

[54] E. A. Yildirim and S. J. Wright, "Warm-start strategies in interior-point methods for linear programming," *SIAM Journal on Optimization*, vol. 12, no. 3, pp. 782–810, 2002.

[55] L. Krishnamachari, D. Estrin, and S. Wicker, "The impact of data aggregation in wireless sensor networks," in *22nd International Conference on Distributed Computing Systems Workshops*. IEEE, 2002, pp. 575–578.

[56] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "TAG: a tiny aggregation service for ad-hoc sensor networks," *ACM SIGOPS Operating Systems Review*, vol. 36, no. SI, pp. 131–146, 2002.

[57] A. Willig, "Recent and emerging topics in wireless industrial communications: a selection," *IEEE Transactions on industrial informatics*, vol. 4, no. 2, pp. 102–124, 2008.

[58] A. Richards and J. P. How, "Robust stable model predictive control with constraint tightening," in *American Control Conference (ACC'06)*, 2006, pp. 1557–1562.

[59] G. C. Goodwin, H. Haimovich, D. E. Quevedo, and J. S. Welsh, "A moving horizon approach to networked control system design," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1427–1445, 2004.

[60] C. V. Rao, J. B. Rawlings, and D. Q. Mayne, "Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations," *IEEE transactions on automatic control*, vol. 48, no. 2, pp. 246–258, 2003.

[61] J. Sijs, M. Lazar, and W. P. M. H. Heemels, "On integration of event-based estimation and robust MPC in a feedback loop," in *13th ACM international conference on Hybrid systems: computation and control (HSCC'10)*. ACM, 2010, pp. 31–40.

[62] D. Bernardini and A. Bemporad, "Energy-aware robust model predictive control based on noisy wireless sensors," *Automatica*, vol. 48, no. 1, pp. 36–44, 2012.

[63] A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos, "The explicit linear quadratic regulator for constrained systems," *Automatica*, vol. 38, no. 1, pp. 3–20, 2002.

[64] D. Lehmann, E. Henriksson, and K. H. Johansson, "Event-triggered model predictive control of discrete-time linear systems subject to disturbances," in *European Control Conference (ECC'13)*. IEEE, 2013, pp. 1156–1161.

[65] D. Antunes and W. P. M. H. Heemels, "Rollout event-triggered control: beyond periodic control performance," *IEEE Transactions on Automatic Control*, vol. 59, no. 12, pp. 3296–3311, 2014.

[66] D. P. Bertsekas, *Dynamic programming and optimal control*, vol. 1, no. 3.

[67] F. D. Brunner, W. P. M. H. Heemels, and F. Allgöwer, "Robust event-triggered MPC with guaranteed asymptotic bound and average sampling rate," *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 5694–5709, 2017.

[68] S. V. Raković, B. Kouvaritakis, R. Findeisen, and M. Cannon, "Homothetic tube model predictive control," *Automatica*, vol. 48, no. 8, pp. 1631–1638, 2012.

[69] B. Demirel, E. Ghadimi, D. E. Quevedo, and M. Johansson, "Optimal control of linear systems with limited control actions: threshold-based event-triggered control," *IEEE Transactions on Control of Network Systems*, vol. (early access), 2017.

[70] J. C. Dunn and D. P. Bertsekas, "Efficient dynamic programming implementations of Newton's method for unconstrained optimal control problems," *Journal of Optimization Theory and Applications*, vol. 63, no. 1, pp. 23–38, 1989.

[71] Y. Wang and S. Boyd, "Fast model predictive control using online optimization," *IEEE Transactions on Control Systems Technology*, vol. 18, no. 2, pp. 267–278, 2010.

[72] S. Richter, C. N. Jones, and M. Morari, "Computational complexity certification for real-time MPC with input constraints based on the fast gradient method," *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1391–1403, 2012.

[73] I. Kolmanovsky and E. G. Gilbert, "Theory and computation of disturbance invariant sets for discrete-time linear systems," *Mathematical problems in engineering*, vol. 4, no. 4, pp. 317–367, 1998.

[74] H. Payne and L. Silverman, "On the discrete time algebraic Riccati equation," *IEEE Transactions on Automatic Control*, vol. 18, no. 3, pp. 226–234, 1973.

[75] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*. SIAM, 1994.

[76] P. Gritzmann and V. Klee, "Computational complexity of inner and outerj-radii of polytopes in finite-dimensional normed spaces," *Mathematical programming*, vol. 59, no. 1-3, pp. 163–213, 1993.

[77] A. Bemporad, C. Filippi, and F. D. Torrisi, "Inner and outer approximations of polytopes using boxes," *Computational Geometry*, vol. 27, no. 2, pp. 151–178, 2004.

[78] A. S. Nemirovsky and D. B. Yudin, *Problem complexity and method efficiency in optimization*. John Wiley & Sons, 1983.

[79] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

[80] J. Renegar, "A polynomial-time algorithm, based on Newton's method, for linear programming," *Mathematical Programming*, vol. 40, no. 1-3, pp. 59–93, 1988.

[81] M. Velasco, J. Fuertes, and P. Marti, "The self triggered task model for real-time control systems," in *Work-in-Progress Session of the 24th IEEE Real-Time Systems Symposium*, vol. 384, 2003.

[82] E. W. Dijkstra, "On the role of scientific thought," in *Selected writings on computing: a personal perspective*. Springer, 1982, pp. 60–66.

[83] G. Buttazzo, G. Lipari, and L. Abeni, "Elastic task model for adaptive rate control," in *Proc. 19th IEEE Real-Time Systems Symposium*, 1998, pp. 286–295.

[84] M. Caccamo, G. Buttazzo, and L. Sha, "Elastic feedback control," in *Proc. 12th Euromicro Conference on Real-Time Systems*, 2000, pp. 121–128.

[85] C. Lu, J. A. Stankovic, S. H. Son, and G. Tao, "Feedback control real-time scheduling: Framework, modeling, and algorithms," *Real-Time Systems*, vol. 23, no. 1-2, pp. 85–126, 2002.

[86] A. Cervin and J. Eker, "Control-scheduling codesign of real-time systems: The control server approach," *Journal of Embedded Computing*, vol. 1, no. 2, pp. 209–224, 2004.

[87] R. Bhattacharya and G. J. Balas, "Anytime control algorithm: Model reduction approach," *Journal of Guidance, Control, and Dynamics*, vol. 27, no. 5, pp. 767–776, 2004.

[88] D. Fontanelli, L. Greco, and A. Bicchi, "Anytime control algorithms for embedded real-time systems," in *Hybrid Systems: Computation and Control*, 2008, pp. 158–171.

[89] S. Al-Areqi, D. Gorges, S. Reimann, and S. Liu, "Event-based control and scheduling codesign of networked embedded control systems," in *American Control Conference (ACC'13)*, 2013, pp. 5299–5304.

[90] S. Al-Areqi, D. Gorges, and S. Liu, "Stochastic event-based control and scheduling of large-scale networked control systems," in *European Control Conference (ECC'14)*, 2014, pp. 2316–2321.

[91] A. Sharifi Kolarijani, D. Adzkiya, and M. Mazo Jr., "Symbolic abstractions for the scheduling of event-triggered control systems," in *54th IEEE Conference on Decision and Control (CDC'15)*, 2015, pp. 6153–6158.

[92] A. S. Kolarijani and M. Mazo Jr., "Formal traffic characterization of lti event-triggered control systems," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 1, pp. 274–283, 2018.

[93] R. Alur and R. Kurshan, "Timing analysis in COSPAN," *Hybrid Systems III*, pp. 220–231, 1996.

[94] T. A. Henzinger, X. Nicollin, J. Sifakis, and S. Yovine, "Symbolic model checking for real-time systems," *Information and computation*, vol. 111, no. 2, pp. 193–244, 1994.

[95] R. Alur and D. L. Dill, "A theory of timed automata," *Theoretical computer science*, vol. 126, no. 2, pp. 183–235, 1994.

[96] K. G. Larsen, P. Pettersson, and W. Yi, "UPPAAL in a nutshell," *International Journal on Software Tools for Technology Transfer*, vol. 1, no. 1, pp. 134–152, 1997.

[97] X. Wang and M. D. Lemmon, "Self-triggered feedback control systems with finite-gain $\mathscr{L}_2$ stability," *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 452–467, 2009.

[98] K. Gu, J. Chen, and V. L. Kharitonov, *Stability of time-delay systems*. Springer Science & Business Media, 2003.

[99] C. Van Loan, "The sensitivity of the matrix exponential," *SIAM Journal on Numerical Analysis*, vol. 14, no. 6, pp. 971–981, 1977.

[100] G. Ewald, *Combinatorial convexity and algebraic geometry*. Springer Science & Business Media, 2012, vol. 168.

[101] P. Tabuada, *Verification and Control of Hybrid Systems: A Symbolic Approach*. Springer, 2009.

[102] C. Fiter, L. Hetel, W. Perruquetti, and J.-P. Richard, "A state dependent sampling for linear state feedback," *Automatica*, vol. 48, no. 8, pp. 1860–1867, 2012.

[103] L. Hetel, J. Daafouz, and C. Iung, "Stabilization of arbitrary switched linear systems with unknown time-varying delays," *IEEE Transactions on Automatic Control*, vol. 51, no. 10, pp. 1668–1674, 2006.

[104] A. Chutinan and B. H. Krogh, "Computing polyhedral approximations to flow pipes for dynamic systems," in *37th IEEE Conference on Decision and Control (CDC'98)*, vol. 2, Dec. 1998, pp. 2089–2094.

[105] M. C. F. Donkers and W. P. M. H. Heemels, "Output-based event-triggered control with guaranteed-gain and improved and decentralized event-triggering," *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1362–1376, 2012.

[106] W. P. M. H. Heemels, M. C. F. Donkers, and A. R. Teel, "Periodic event-triggered control for linear systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 847–861, 2013.