



HAL
open science

Near-optimal control of discrete-time nonlinear systems with stability guarantees

Mathieu Granzotto

► **To cite this version:**

Mathieu Granzotto. Near-optimal control of discrete-time nonlinear systems with stability guarantees. Automatic. Université de Lorraine, 2019. English. NNT : 2019LORR0301 . tel-02517304

HAL Id: tel-02517304

<https://hal.univ-lorraine.fr/tel-02517304>

Submitted on 24 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Near-optimal control of discrete-time nonlinear systems with stability guarantees

THÈSE

présentée et soutenue publiquement le 18 décembre 2019

pour l'obtention du

Doctorat de l'Université de Lorraine

(mention Automatique)

par

Mathieu Granzotto

Composition du jury

<i>Rapporteurs :</i>	Paulo Tabuada Emmanuel Trélat	Professeur à UCLA, États-unis Professeur à Sorbonne Université
<i>Examineurs :</i>	Cristina Stoica Maniu Didier Henrion Dragan Nešić	Professeur à CentraleSupélec Directeur de Recherche CNRS, LAAS Professeur à University of Melbourne, Australie
<i>Invité :</i>	Bruno Scherrer	Chargé de recherche INRIA, IECL
<i>Encadrants :</i>	Jamal Daafouz Romain Postoyan	Professeur à Université de Lorraine Chargé de recherche CNRS, CRAN

Mis en page avec la classe thesul.

Acknowledgments

There are moments in life that we are obliged to wonder how we got here. For me, this is one of these moments. Four years ago, I would never have expected to be flung in such a captivating and delightful atmosphere that is the CRAN. I have many to thank it for. Firstly, my (some unofficially) advisors, Romain Postoyan, Jamal Daafouz, Lucian Buşoniu and Dragan Nešić, whose valuable guidance and support meet no equal. I want to specially single out Romain, as his dedication and knowledge serves everyday as inspiration to provide a better work, a stronger result, a clearer argument. Under his unwavering supervision, I am certain that I became a better scientist and, more importantly, a better person. Secondly, I have to thank the several PhD candidates and young researchers that I've met at CRAN. To name them all would be a daunting task, but the fruitful and wide-ranging discussion have made these last years lighter. Be either the right timing for pasta, multiple economical models, the quirks of our many languages, cultures and origins, our political future, of academics, of climate change, or the games to play, computer hardware and what-else, discussing with you all made my day. I hope we will keep in touch.

There are also the ones from before my thesis that have nonetheless enriched my understanding and my life. Friends from the US, from about everywhere from Brasil, India and China. Many of those that I had, and some that I still will, aggravate by needlessly messaging far-fetched tangents and bad jokes.

I am also very grateful for my family, in particular my parents, which have gave me a lifetime of opportunity, of education and unconditional support. Also my grandparents, whose stories have guided my life. Finally, but not least, I want to thank my Fiancée Julia, whose support and spirit always push me to higher challenges and to take every opportunity to its fullest. There is no better person to share the adventures and soul-enriching experiences that this world provides, and the calmness of the day-to-day life.

To the trees, for without papers, research would not endure.

Contents

Notation	xi
Chapter 1 Introduction	1
1.1 Optimal control of discrete-time systems	1
1.2 Linear quadratic problem	2
1.3 Dynamic programming	4
1.4 Approximate value iteration	7
1.5 Discount factor	8
1.6 What about stability?	9
1.7 Value iteration versus model predictive control	10
1.8 Exploiting stability	11
1.9 Organization and summary of the contributions	12
Chapter 2 Finite-horizon discounted optimal control	15
2.1 Introduction	15
2.2 Problem statement	16
2.3 Stability results	20
2.3.1 Lyapunov properties	20
2.3.2 Main result	24
2.3.3 Stronger stability properties	29
2.3.4 Comparison of the conditions on γ and d with existing results	33
2.4 Relationships between optimal value functions	35
2.5 Near-optimal sequence of inputs	39
2.6 Illustrative example: an inverted pendulum	47
2.7 Conclusion	50

Chapter 3 Control of switched nonlinear discrete-time systems: a planning-based approach	53
3.1 Introduction	53
3.2 Problem statement	54
3.3 OP_{\min}	55
3.3.1 Main idea	55
3.3.2 Assumptions	58
3.3.3 Algorithm description	60
3.4 Main results	65
3.4.1 Relationship between OP_{\min} and $V_{\infty}(x)$	65
3.4.2 Stability	67
3.4.3 Nominal robustness	75
3.4.4 Near-optimality guarantees	75
3.5 Example: a cubic integrator	82
3.6 Case study	86
3.6.1 Main result	86
3.6.2 Example: a flexible joint robotic arm	89
3.7 Conclusion	92
Chapter 4 A regret perspective	95
4.1 Introduction	95
4.2 The notion of regret	96
4.2.1 Definition	96
4.2.2 Near-optimality guarantees	99
4.2.3 Stability guarantees	102
4.3 Examples of near-optimal control	103
4.3.1 An unstable optimal feedback law	103
4.3.2 Finite-horizon optimal feedback law	104
4.4 Conclusion	105
Chapter 5 Conclusions	107
5.1 Overview	107
5.2 Perspectives	109
5.2.1 Short term	109
5.2.2 Longer term	110

Appendix A Technical results for Chapter 2	113
Appendix B Technical results for Chapter 3	115
B.1 Auxiliary Result for Theorem 3.4	115
B.2 Semiglobal asymptotic stability property for $OP_{\min, \text{stop}}$	116
B.3 Average cost and cost-to-attractor guarantees for $OP_{\min, \text{stop}}$	118
Appendix C Additional simulation plots for Chapter 3	123
Appendix D Stability guarantees for nonlinear discrete-time systems controlled by approximate value iteration	127
Bibliography	139

List of Figures

2.1	$\sigma(\hat{\phi}(\cdot, x))$ for 4 different pairs (γ, d) and $x = (3, 0)$	49
2.2	Convergence testing of 1000 sample pairs (γ, d) , $N = 33^2$. Symbol \circ denotes convergence to \mathcal{S} , while \times the converse.	49
2.3	Convergence testing of 1000 sample pairs (γ, d) , $N = 31^2$. Symbol \circ denotes convergence to \mathcal{S} , while \times the converse.	50
3.1	State evolution for $\text{OP}_{\text{min,budget}}$ with $B = 3000$ and $x = (-1, 1.5)$	83
3.2	$\sigma(\phi(\cdot, x))$ for inputs given by $\text{OP}_{\text{min,stop}}$, $g(1, x)$ and $g(2, x)$	91
3.3	Input selection of $\text{OP}_{\text{min,stop}}$ for $x_0 = (10, -10, -10, -10)$	91
C.1	State evolution for $\text{OP}_{\text{min,stop}}$ with $\varepsilon = 5$ and $x = (10, 15)$	124
C.2	State evolution for $\text{OP}_{\text{min,stop}}$ with $\varepsilon = \frac{10}{9} = 1.111$ and $x = (10, 15)$	124
C.3	State evolution for $\text{OP}_{\text{min,stop}}$ with $\varepsilon = \frac{10}{11} = 0.909$ and $x = (10, 15)$	125
C.4	State evolution for $\text{OP}_{\text{min,stop}}$ with $\varepsilon = \frac{10}{12} = 0.833$ and $x = (10, 15)$	125
C.5	State evolution for $\text{OP}_{\text{min,stop}}$ with $\varepsilon = \frac{10}{17} = 0.588$ and $x = (10, 15)$	126
C.6	State evolution for $\text{OP}_{\text{min,stop}}$ with $\varepsilon = \frac{1}{9} = 0.111$ and $x = (10, 15)$	126

List of Figures

Notation

\mathbb{R}	Set of real numbers, i.e. $(-\infty, \infty)$
$\mathbb{R}_{\geq 0}$	Set of positive real numbers, i.e. $[0, \infty)$
$\mathbb{R}_{> 0}$	Set of nonnegative real numbers, i.e. $(0, \infty)$
$\mathbb{Z}_{\geq 0}$	Set of positive integers, i.e. $\{0, 1, 2, \dots\}$
$\mathbb{Z}_{> 0}$	Set of nonnegative integers, i.e. $\{1, 2, \dots\}$
(x, y)	$[x^\top, y^\top]^\top$, with $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ and $n, m \in \mathbb{Z}_{> 0}$
$\chi \in \mathcal{K}$	Continuous function $\chi : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$, zero at zero, strictly increasing
$\chi \in \mathcal{K}_\infty$	$\chi \in \mathcal{K}$ and unbounded
$\beta \in \mathcal{KL}$	Continuous function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$, $\beta(\cdot, t) \in \mathcal{K}$ for any $t \geq 0$, $\beta(s, \cdot)$ is decreasing to 0 for any $s \geq 0$
\mathbb{I}	Identity map from $\mathbb{R}_{\geq 0}$ to $\mathbb{R}_{\geq 0}$
\mathbf{u}	A sequence of inputs $[u_0, u_1, \dots]$ of length $d \in \mathbb{Z}_{\geq 0} \cup \{\infty\}$ $u_i \in \mathbb{R}^m$ for $i \in \{0, \dots, d\}$
$\mathbf{u} _k$	The truncation of sequence \mathbf{u} to its first k elements, i.e. $\mathbf{u} _k = [u_0, \dots, u_{k-1}]$ $\mathbf{u} _0 = \emptyset$ by convention.
$f^{(k)}$	The composition of f to itself k times, i.e. for $f : \mathbb{R} \rightarrow \mathbb{R}$, and $k \in \mathbb{Z}_{\geq 0}$, $f^{(k)} = f(f^{(k-1)})$ $f^{(0)} = \mathbb{I}$.
$\lfloor \cdot \rfloor$	Floor function
$ x $	The Euclidean norm of a vector $x \in \mathbb{R}^n$
$ x _{\mathcal{A}}$	The distance of a vector $x \in \mathbb{R}^n$ to non-empty set \mathcal{A} defined as $ x _{\mathcal{A}} = \inf\{ z - x : z \in \mathcal{A}\}$
$\lambda_{\min}(P)$	the smallest eigenvalue of a symmetric real matrix P
$\lambda_{\max}(P)$	the largest eigenvalue of a symmetric real matrix P

Glossary

DP	Dynamic Programming
ADP	Approximate Dynamic Programming
VI	Value Iteration
AVI	Approximated Value Iteration
PI	Policy Iteration
API	Approximated Policy Iteration
OP	Optimistic Planning
RL	Reinforcement Learning
MPC	Model Predictive Control
NMPC	Nonlinear Model Predictive Control
LQR	Linear Quadratic Regulator

Chapter 1. Introduction

1.1	Optimal control of discrete-time systems	1
1.2	Linear quadratic problem	2
1.3	Dynamic programming	4
1.4	Approximate value iteration	7
1.5	Discount factor	8
1.6	What about stability?	9
1.7	Value iteration versus model predictive control	10
1.8	Exploiting stability	11
1.9	Organization and summary of the contributions	12

1.1 Optimal control of discrete-time systems

Optimal control consists in designing the inputs to a dynamical system so that these optimize a given cost function. Its applications are wide ranging from path planning of spacecrafts [61], increased flight endurance of drones [51], to better performing switching power converters [19] to cite only a few examples. Optimal control inception can be traced back from the Brachistochrone curve problem [68]. Proposed in 1696 by J. Bernoulli, the challenge was to find “the curve traced out by a point acted only by gravity, which starts at A and reaches B in the shortest time”. Solved by several mathematicians of that time, this eventually lead to the birth of the calculus of variations, and finally, towards

Pontryagin's Maximum Principle. It is accepted that modern optimal control flourished after the formulation of Pontryagin's Maximum Principle in the 1950s, and numerous tools and applications followed, see e.g. [68].

In this thesis, we concentrate on the optimal control of dynamical systems evolving in discrete time. In this context, we consider the system

$$x^+ = f(x, u), \quad (1.1)$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathcal{U}(x) \subseteq \mathbb{R}^m$ is the input, $f : \mathcal{W} \rightarrow \mathbb{R}^n$ is the deterministic jump map where $\mathcal{W} := \{(x, u) : x \in \mathbb{R}^n, u \in \mathcal{U}(x)\}$. We denote $\mathcal{U}(x)$ the nonempty set of *admissible inputs* for state x . For example, $\mathcal{U}(x)$ can be: the whole set \mathbb{R}^m ; a given compact set, which is relevant to model input saturation; a discrete subset of $\mathbb{Z}_{>0}$, as when inputs are quantized, or when u represents a switching signal [43]; the set of linear feedback controllers. We use $\phi(k, x, \mathbf{u}|_k)$ to denote the solution to system (1.1) at time $k \in \mathbb{Z}_{\geq 0}$ with initial condition x and inputs $\mathbf{u}|_k = [u_0, u_1, \dots, u_{k-1}]$, with the convention $\phi(0, x, \cdot) = \phi(0, x, \emptyset) = x$.

In general, the cost function takes the following form

$$J_\infty(i, x, \mathbf{u}) := \sum_{k=i}^{\infty} \ell(k, \phi(k, x, \mathbf{u}|_k), u_k) \quad (1.2)$$

where $\ell : \mathbb{Z}_{\geq 0} \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ is a time-varying *stage cost* and $i \in \mathbb{Z}_{\geq 0}$ is the initial time-step. The objective is to find a sequence of admissible inputs $u \in \mathcal{U}(x)$ to minimize (1.2), that is^{†1}

$$V_\infty(i, x) := \min_{\mathbf{u}} J_\infty(i, x, \mathbf{u}), \quad (1.3)$$

and V_∞ is called the *optimal value function*, which is the minimum value of cost (1.2) for the considered admissible inputs.

1.2 Linear quadratic problem

The problem formulated above is notoriously difficult in general. However, there is an important case for which simple and elegant solutions exist: linear systems and quadratic

^{†1} Each time a minimization problem is formulated over inputs or sequences of inputs, it is implicitly assumed that this is done over admissible inputs.

costs. Indeed, consider the case where system (1.1) is given by

$$x^+ = Ax + Bu \quad (1.4)$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $x \in \mathbb{R}^n$ and $u \in \mathcal{U}(x) = \mathbb{R}^m$, and the cost in (1.2) takes the form

$$J_\infty(x, \mathbf{u}) := \sum_{k=0}^{\infty} \phi(k, x, \mathbf{u}|_k)^\top Q \phi(k, x, \mathbf{u}|_k) + u_k^\top R u_k, \quad (1.5)$$

where $Q \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{m \times m}$. This is known as the Linear Quadratic Regulator (LQR) problem. R.E. Kalman [38] famously solved the LQR problem in continuous-time^{†2} when Q and R are symmetric, positive definite matrices, given that pair (A, B) is controllable, a notion he introduced himself. The solution of the LQR problem for system (1.4) is given by

$$V_\infty(x) := \min_{\mathbf{u}} J_\infty(x, \mathbf{u}) = x^\top P x, \quad (1.6)$$

where $P \in \mathbb{R}^{n \times n}$ is the unique solution of the discrete Riccati equation

$$P = A^\top P A - A^\top P B (R + B^\top P B)^{-1} B^\top P A + Q, \quad (1.7)$$

see [2]. Moreover, the optimal inputs are generated by the uniquely defined linear state-feedback controller

$$K := (B^\top P B + R)^{-1} B^\top P A, \quad (1.8)$$

i.e. $J_\infty(x, \mathbf{u}_\infty^*(x)) = x^\top P x$, where $\mathbf{u}_\infty^*(x) = [u_0^*, u_1^*, \dots]$ and $u_k^* = -K \phi(k, x, \mathbf{u}_\infty^*(x)|_k)$. For more details on the LQ control, the reader is invited to consult, e.g., [2, 7, 38].

Controller (1.8) thus ensures optimality. Additional assumptions need to be made for the closed-loop system (1.4), (1.8) to exhibit stability properties, as noted by R.E. Kalman in [38]: “In the engineering literature it is often assumed (tacitly and incorrectly) that a system with optimal control law is necessarily stable”. The extra requirement is a detectability condition, as formalized next.

Proposition 1.1. *If (A, C) is observable where C is a real matrix such that $Q = C^\top C$, the origin of system $x^+ = (A - BK)x$ with K defined in (1.8) is globally exponentially*

^{†2} He actually solved the linear time-varying case, i.e. matrices A, B, Q, R are free to vary in time, which is more general.

stable, i.e. there exist $\lambda_1 \geq 1$ and $\lambda_2 > 0$ such that for any initial condition $x \in \mathbb{R}^n$, $|\phi(k, x, \mathbf{u}_\infty^*(x))| \leq \lambda_1 e^{-\lambda_2 k} |x|$. \square

We refer to Proposition 4.4.1 in [8] for the proof that $A - BK$ is Schur, i.e. its eigenvalues lie in the open unit disk of \mathbb{C} , which implies in global exponential stability.

1.3 Dynamic programming

Going back to system (1.1) and cost (1.2), it is very difficult, if not impossible, to construct explicitly an optimal sequence of inputs in general when the system is nonlinear and the stage cost is not quadratic. Indeed, many relevant choices of stage costs ℓ exist beside the quadratic one in (1.5). For instance, when we are interested in making the solutions to (1.1) converge to a given set or when we are only interested in minimizing certain variables, we can be tempted to define $\ell(x, u)$ as $|x|_{\mathcal{A}} + \lambda(u)$ with $\lambda(u) \geq 0$ and a given set $\mathcal{A} \subset \mathbb{R}^n$. In the Reinforcement Learning (RL) literature, costs of the following form are also frequently encountered [67]

$$\ell(k, x, u) = \begin{cases} 1 & \text{when } x \notin \mathcal{A} \\ 0 & \text{when } x \in \mathcal{A}, \end{cases} \quad (1.9)$$

for any $k \in \mathbb{Z}_{\geq 0}$, u in some compact subset of \mathbb{R}^m and non-empty set $\mathcal{A} \subset \mathbb{R}^n$. The underlying idea is to penalize the state when it is not in a desired set \mathcal{A} and manifest clearly our desire to reach (and stay in) \mathcal{A} in minimal time.

As we already mentioned, considering general nonlinear systems and stage costs is appealing, but it is difficult to analytically construct optimal inputs. A powerful approach to overcome this obstacle is to resort to Dynamic Programming (DP)^{†3}. Popularized by R.E. Bellman, DP revolves around a recursive property for the optimal cost (1.3), when it exists, which is, for any $k \in \mathbb{Z}_{\geq 0}$ and $x \in \mathbb{R}^n$,

$$V_\infty(k, x) = \min_u \left\{ \ell(k, x, u) + V_\infty(k + 1, f(x, u)) \right\}, \quad (1.10)$$

^{†3} Alternative approaches exist for specific classes of systems. For instance, occupation measure and moment relaxations are a possible approach for polynomial systems with polynomial stage costs [64], see [41] for results for continuous-time systems.

also called *Bellman equation* [7]. As R.E. Bellman wrote in [5]: “An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision”, which is captured by (1.10). Instead of optimizing an infinite sequence of inputs \mathbf{u} , equation (1.10) collapses the optimization problem into a single decision u . In view of (1.10), the optimal inputs are thus given by

$$u_0^*(k, x) \in \arg \min_{u \in \mathcal{U}(x)} \left\{ \ell(k, x, u) + V_\infty(k+1, f(x, u)) \right\}, \quad (1.11)$$

which may be non-unique, justifying the usage of the symbol “ \in ” in (1.11). Equation (1.11) is appealing but there is a catch: we need to know V_∞ in order to construct u_0^* . It appears that to know V_∞ for any $k \in \mathbb{Z}_{\geq 0}$ and $x \in \mathbb{R}^n$ is very difficult in general. Another issue is that representing V_∞ numerically might be too costly, if not intractable. This is particularly an issue for higher-dimensional systems, due to the so-called *curse of dimensionality*. An alternative consists in looking for approximations of V_∞ instead, based on which near-optimal inputs are constructed; this corresponds to *Approximate Dynamic Programming* (ADP). To explain how this can be done, consider *discounted* stage costs, i.e. $\ell(k, x, u) = \gamma^k \ell(x, u)$ with some abuse of notation, where $\gamma \in (0, 1]$ is a possible discount factor, and $\ell(x, u) \geq 0$ for any $k \in \mathbb{Z}_{\geq 0}$ and $(x, u) \in \mathcal{W}$. We explain in the following why it is useful to introduce a discount factor here, see Section 1.5. In sum, we want to minimize, for any $x \in \mathbb{R}^n$,

$$J_{\gamma, \infty}(x, \mathbf{u}) := \sum_{k=0}^{\infty} \gamma^k \ell(\phi(k, x, \mathbf{u}|_k), u_k). \quad (1.12)$$

Bellman equation for (1.12) is then given by

$$V_{\gamma, \infty}(x) = \min_u \left\{ \ell(x, u) + \gamma V_{\gamma, \infty}(f(x, u)) \right\}. \quad (1.13)$$

Value Iteration (VI) [7] is an iterative procedure based on (1.13) which we now describe. Given an initial cost function $V_{\gamma, -1}$, VI generates a sequence of value functions $V_{\gamma, d}$ for any $d \in \mathbb{Z}_{\geq 0}$ as follows, for any $x \in \mathbb{R}^n$,

$$V_{\gamma, d}(x) = \min_u \left\{ \ell(x, u) + \gamma V_{\gamma, d-1}(f(x, u)) \right\}. \quad (1.14)$$

Note the difference with (1.13) where $V_{\gamma,\infty}$ appears on both sides of the equality, while here in (1.14) the value at the next iteration, $V_{\gamma,d}$, is defined based on the value of the previous iteration, i.e. $V_{\gamma,d-1}$. Given $d \in \mathbb{Z}_{\geq 0}$, the associated input, also called policy, is defined as

$$u_{\gamma,d}^*(x) \in \arg \min_u \left\{ \ell(x, u) + \gamma V_{\gamma,d-1}(f(x, u)) \right\}. \quad (1.15)$$

It can be shown that for $d \rightarrow \infty$, $V_{\gamma,d}$ converges to the fixed-point of (1.14), i.e. $V_{\gamma,\infty}$, when it exists, see e.g. Proposition 1.6.4 from [7] when $\gamma \in (0, 1)$, and [9] when $\gamma = 1$, under some conditions; see also [1, 27, 44, 48]. We state this more precisely in the following lemma when stage cost ℓ is such that there exists $M \geq 0$ such that $\ell(x, u) \in [0, M]$ for any $(x, u) \in \mathcal{W}$, inspired by the arguments in Proposition 1.2.1 from [7] when $\gamma \in (0, 1)$.

Lemma 1.1. *For any $x \in \mathbb{R}^n$, $\gamma \in (0, 1)$, let the initial value function $V_{\gamma,-1}$ be such that $0 \leq V_{\gamma,-1}(x) \leq V_{\gamma,\infty}(x)$ and consider the case where there exists $M \geq 0$ such that $\ell(x, u) \in [0, M]$ for any $(x, u) \in \mathcal{W}$. Then the following holds for any $d \in \mathbb{Z}_{\geq 0}$,*

$$0 \leq V_{\gamma,\infty}(x) - V_{\gamma,d}(x) \leq \gamma^{d+1} \frac{1}{1-\gamma} M. \quad (1.16)$$

□

Proof. Let $\gamma \in (0, 1)$. We first show that $V_{\gamma,d}(x) \leq V_{\gamma,\infty}(x)$ for any $x \in \mathbb{R}^n$ and any $d \in \mathbb{Z}_{\geq 0}$, and proceed by contradiction. Assume there exists $x \in \mathbb{R}^n$ such that $V_{\gamma,\infty}(x) < V_{\gamma,0}(x)$. On the other hand, in view of (1.14),

$$V_{\gamma,0}(x) = \min_u \left\{ \ell(x, u) + \gamma V_{\gamma,-1}(f(x, u)) \right\} \leq \min_u \left\{ \ell(x, u) + \gamma V_{\gamma,\infty}(f(x, u)) \right\}, \quad (1.17)$$

as $V_{\gamma,-1} \leq V_{\gamma,\infty}$. In view of (1.13), we derive $V_{\gamma,\infty}(x) < V_{\gamma,0}(x) \leq \min_u \left\{ \ell(x, u) + \gamma V_{\gamma,\infty}(f(x, u)) \right\} = V_{\gamma,\infty}(x)$, a contradiction is attained. By proceeding iteratively, one concludes that $V_{\gamma,d}(x) \leq V_{\gamma,\infty}(x)$ for any $x \in \mathbb{R}^n$ and any $d \in \mathbb{Z}_{\geq 0}$. Therefore, the lower-bound in (1.16) holds.

In view of (1.14) and (1.15), there is a sequence $\mathbf{u}_0 = [u_{\gamma,0}^*(x)]$ such that $V_{\gamma,0}(x) - \gamma V_{\gamma,-1}(\phi(1, x, \mathbf{u}_0|_1)) = J_{\gamma,0}(x, \mathbf{u}_0)$. Again, by proceeding by iteration, we conclude that there exist a sequence $\mathbf{u}_d = [u_0, \dots, u_d]$ such that $V_{\gamma,d}(x) - \gamma^{d+1} V_{\gamma,-1}(\phi(d +$

$1, x, \mathbf{u}|_{d+1}) = J_{\gamma,d}(x, \mathbf{u}_d)$, and since $V_{\gamma,-1} \geq 0$, $V_{\gamma,d}(x) \geq J_{\gamma,d}(x, \mathbf{u}_d)$. Consider sequence $\hat{\mathbf{u}} = [\mathbf{u}_d, \mathbf{u}_{\gamma,\infty}^*(\phi(d+1, x, \mathbf{u}_d|_{d+1}))]$, which consists of the inputs \mathbf{u}_d followed by an optimal input sequence of infinite length at state $\phi(d+1, x, \mathbf{u}_d|_{d+1})$. We derive $J_{\gamma,\infty}(x, \hat{\mathbf{u}}) = J_{\gamma,d}(x, \mathbf{u}_d) + \gamma^{d+1}J_{\gamma,\infty}(\phi(d+1, x, \mathbf{u}_d|_{d+1}), \mathbf{u}_{\gamma,\infty}^*(\phi(d+1, x, \mathbf{u}_d|_{d+1})))$. Since $\mathbf{u}_{\gamma,\infty}^*(\phi(d+1, x, \mathbf{u}_d|_{d+1}))$ is the optimal sequence for state $\phi(d+1, x, \mathbf{u}_d|_{d+1})$ and $J_{\gamma,d}(x, \mathbf{u}_d) \leq V_{\gamma,d}(x)$, it follows that

$$J_{\gamma,\infty}(x, \hat{\mathbf{u}}) \leq V_{\gamma,d}(x) + \gamma^{d+1}V_{\gamma,\infty}(\phi(d+1, x, \mathbf{u}_d|_{d+1})). \quad (1.18)$$

On the other hand, since $V_{\gamma,\infty}(x)$ is the optimal cost, any other infinite-length sequence will have a higher cost, i.e. $V_{\gamma,\infty}(x) \leq J_{\gamma,\infty}(x, \hat{\mathbf{u}})$. Therefore, in view of (1.18),

$$V_{\gamma,\infty}(x) \leq V_{\gamma,d}(x) + \gamma^{d+1}V_{\gamma,\infty}(\phi(d+1, x, \mathbf{u}_d|_{d+1})). \quad (1.19)$$

Since ℓ lives in $[0, M]$, $\sum_{k=0}^{\infty} \gamma^k \ell(\phi(k, x, \mathbf{u}), u_k) \leq M \sum_{k=0}^{\infty} \gamma^k$ for any admissible sequence $\mathbf{u} = [u_0, u_1, \dots]$ and $x \in \mathbb{R}^n$, thus $V_{\gamma,\infty}(x) \leq V_{\gamma,d}(x) + \frac{\gamma^{d+1}}{1-\gamma}M$ and (1.16) holds. \blacksquare

Lemma 1.1 provides near-optimality guarantees for value iteration, as $\frac{\gamma^{d+1}}{1-\gamma}M$ is an upperbound on the error between the optimal value function at x , $V_{\gamma,\infty}(x)$, and the estimated value obtained by VI at step d , $V_{\gamma,d}(x)$. We see that when d tends to infinity, this bound goes to zero: $V_{\gamma,d}(x)$ converges to $V_{\gamma,\infty}(x)$.

Remark 1.1. Another general approach to generate near-optimal inputs is Policy Iteration (PI), see [7]. In this thesis, we concentrate on VI or similar algorithms. However, we believe that the presented results can be extended to (Approximate) Policy Iteration ((A)PI) under mild changes. This is an interesting and relevant future work direction, as mentioned in Chapter 5. \square

1.4 Approximate value iteration

It appears that to implement VI exactly as described in (1.14) and (1.15) is virtually impossible in particular when the state space is continuous, which is the case here. In practice, only approximate versions of VI are applied, we talk about Approximate Value

Iteration (AVI). To convey the main idea, we consider the simplest approach where we constrain the state space to be a compact set $\mathcal{X} \subset \mathbb{R}^n$, which we quantize in finite number of regions. The idea is to assign a single constant value to the discounted value function at each iteration, $V_{\gamma,d}$. We thus generate an approximation of $V_{\gamma,d}$ in (1.14) which is piece-wise constant. Under mild conditions, this approximation can be made as close as desired to $V_{\gamma,d}$ by refining the quantization of \mathcal{X} . Various other schemes exist in the literature, see e.g. [10, 16]. In all cases, the approximation of $V_{\gamma,d}$, denoted by $\widehat{V}_{\gamma,d}$ is such that, for any $x \in \mathcal{X}$ and $\gamma \in (0, 1)$ and $d \in \mathbb{Z}_{\geq 0}$,

$$\widehat{V}_{\gamma,d}(x) = \min_u \left\{ \ell(x, u) + \gamma \widehat{V}_{\gamma,d-1}(f(x, u)) \right\} + \epsilon_{\gamma,d}(x), \quad (1.20)$$

where $\epsilon_{\gamma,d}$ is due to the approximation errors, which typically can be made as small as desired by tuning the approximation scheme. As a result, when $\epsilon_{\gamma,d}$ is constant, recurrent error bounds between $\widehat{V}_{\gamma,d}$ and $V_{\gamma,\infty}$ are given by [52], for any $x \in \mathbb{R}^n$,

$$|\widehat{V}_{\gamma,d}(x) - V_{\gamma,\infty}(x)| \leq \frac{\gamma^d}{1-\gamma} \epsilon_1 + \epsilon_2, \quad (1.21)$$

where $\epsilon_1, \epsilon_2 > 0$ are errors associated to the approximation. See for more details, e.g. [10, 28, 46, 47, 52, 65, 66].

1.5 Discount factor

Let us go back on the choice of considering a discount factor γ in the presentation of (A)VI above. In one hand, the error bound from Lemma 1.1 vanishes for d large, since $\gamma^d \rightarrow 0$ when $d \rightarrow \infty$ and $\gamma \in (0, 1)$. In the other hand, when $\gamma \rightarrow 1$, the last bound in (1.16) explodes to infinity. In this case, extra conditions need to be made to ensure the convergence of VI, see [9] for instance as already mentioned. On the other hand, notice that, in general, $V_{\gamma,\infty}$ finite only implies that $\gamma^k \ell(\phi(k, x, \mathbf{u}_\infty|_k), \mathbf{u}_k^*) \rightarrow 0$ as k tends to ∞ : the stage cost $\ell(\phi(k, x, \mathbf{u}_\infty|_k), \mathbf{u}_k^*)$ may diverge. If the given ℓ is quadratic for instance, like in Section 1.2, the obtained closed-loop system might be unstable. The discount factor poses a dilemma between guaranteeing (fast) convergence of VI to $V_{\gamma,\infty}$ and stability. Discount factors are also important in ADP to cope with approximation errors [10], as the

bound in (1.21) also explodes when $\gamma \rightarrow 1$. The underlying reason of these properties is that $\gamma \in (0, 1)$ makes Bellman operator contractive, see [7].

1.6 What about stability?

While VI can be used to generate near-optimal inputs for general systems (1.1) and (discounted) stage costs, it remains unclear whether the obtained closed-loop system ensures stability guarantees. This is a fundamental question as stability is critical in many applications, as it provides analytical guarantees on the behaviour of the controlled system solutions as time evolves, whether they converge to a desired point or set, and whether if they are initially closed to an attractor, they will remain so for all future time. Stability also ensures nominal robustness to unmodelled disturbances, although special care is required to ensure this for nonlinear discrete-time systems [40], see also Chapters 2.3.2 and 3.4.3.

In [30, 75], stability results for VI can be found assuming the knowledge of an initial globally stabilizing controller which may be difficult to determine. The author of [31] relaxes this condition. Nevertheless, these works exhibit the next limitations. First, only the stability of a point is investigated, while the closed-loop system may have a more general type of attractors. For example, we might want the attractor to be a non-singleton compact set, like when we want to regulate the temperature of a room to a given range and not necessarily to a precise value. We might also want to stabilize a part of the state variable, loosely speaking, as in partial stability [71], like for time-varying systems where the time is considered as a state variable. Second, we want to consider more general stage costs, and not necessarily quadratic, nor stage costs of the form $Q(x) + R(u)$ with Q and R positive definite functions, and instead accept costs like (1.9). Third, these results do not cope with discounted costs. Fourth, the assumptions made in [30, 75] are subject to some conservatism as they do ensure asymptotic stability properties while we know, from the parallel between VI and Model Predictive Control (MPC) [6] (see also Section 1.7), that, in general, only practical stability properties can be achieved for nonlinear systems [24], in the sense that solutions do not a priori asymptotically converge to the attractor,

but to any desired neighborhood of it by tuning the number of iterations d .

Stability results for systems controlled by AVI can also be found in [31, 46, 74]. In addition to the limitations pointed out above, they rely on strong assumptions on the approximation errors. In particular, the results in [46, 74] make assumptions related to the optimal value function, which we do not know and may thus be difficult to verify. In [31] on the other hand, the approximation errors are required to vanish in the attractor, which is very difficult to ensure.

There is therefore a need for generic tools to analyse the stability of nonlinear discrete-time systems controlled by (A)VI. This is the main objective of this thesis.

Remark 1.2. Stability results for nonlinear discrete-time systems controlled by PI can be found in [13, 29, 45]. The reader interested in results for continuous-time systems can refer to, e.g., [11, 12, 35, 36, 53, 70, 72, 73] \square

1.7 Value iteration versus model predictive control

It turns out that a clear link between VI and MPC can be made [6]. Given $\gamma \in (0, 1]$, suppose we initialize VI with $V_{\gamma,-1}(x) = 0$ for every $x \in \mathbb{R}^n$. The first iteration of VI is, for any $x \in \mathbb{R}^n$, given by

$$V_{\gamma,0}(x) = \min_u \left\{ \ell(x, u) + \gamma V_{\gamma,-1}(f(x, u)) \right\} = \min_u \ell(x, u). \quad (1.22)$$

Similarly, for the second iteration,

$$\begin{aligned} V_{\gamma,1}(x) &= \min_u \left\{ \ell(x, u) + \gamma V_{\gamma,0}(f(x, u)) \right\} \\ &= \min_u \left\{ \ell(x, u) + \gamma \min_{u'} \ell(f(x, u), u') \right\} \\ &= \min_{u, u'} \left\{ \ell(x, u) + \gamma \ell(f(x, u), u') \right\} \\ &= \min_u \sum_{k=0}^1 \gamma^k \ell(\phi(k, x, \mathbf{u}_\infty|_k), u_k^*). \end{aligned} \quad (1.23)$$

Therefore, for any iteration $d \in \mathbb{Z}_{\geq 0}$,

$$V_{\gamma,d}(x) = \min_u \sum_{k=0}^d \gamma^k \ell(\phi(k, x, \mathbf{u}_\infty|_k), u_k^*). \quad (1.24)$$

We observe that VI is in fact solving a finite-horizon discounted optimal control problem, akin to those found in the MPC literature [49]. Also, similar to MPC, inputs computed by (1.15) naturally implies in a receding-horizon controller. Indeed, (1.15) corresponds to running VI to generate $V_{\gamma,d}$ and an associated input, and repeating the process at the next time-step. This is good news a priori, as we could in principle resort to MPC results for cost function with no terminal cost^{†4} to assess the stability of systems controlled by VI. The difficulty is that there is a discount factor in (1.12). We are not aware of any stability results for such costs. In [55], results can be found on the stability of nonlinear discrete-time systems controlled by an optimal sequence of inputs for *infinite-horizon* discounted cost. However, when the sequence of inputs is only *near-optimal*, like for VI, the existing results in [55] requires the algorithm near-optimality bounds, i.e. a bound on the mismatch between $V_{\gamma,\infty}$ and $V_{\gamma,d}$, to be uniform with respect to γ , which is not the case in general, as we saw for instance in Lemma 1.1. On the other hand, in [24] for instance, generic conditions are proposed when the cost is finite-horizon but undiscounted. However this result does not adapt “off-the-shelf” to the discounted case. We thus propose to combine these two approaches, namely [24] and [55], to investigate the stability of system (1.1) controlled by (A)VI. This is very technically challenging as we will see in Chapter 2.

1.8 Exploiting stability

Once stability for the considered controlled system is established, a tempting question is whether we can exploit stability guarantees to our benefit. The answer is yes and this can be done in several ways.

Improved near-optimality bounds. Stability of the closed-loop system provides nice

^{†4} It can be shown that VI initialized to $V_{\gamma,-1} \neq 0$ is equivalent to MPC with terminal costs, i.e.

$$V_{\gamma,d}(x) = \min_{\mathbf{u}} \left\{ \sum_{k=0}^d \gamma^k \ell(\phi(k, x, \mathbf{u}_{\infty}|_k), u_k^*) + \gamma^{d+1} V_{\gamma,-1}(\phi(d+1, x, \mathbf{u}_{\infty}|_{d+1})) \right\}. \quad (1.25)$$

While in this manuscript we study the case without terminal costs, tackling (1.25) is an interesting and relevant future work direction, as mentioned in Chapter 5.

convergence properties of the stage cost along solutions. This can be exploited to revisit near-optimality analysis, i.e. the mismatch between $V_{\gamma,\infty}$ and $V_{\gamma,d}$, as done in Lemma 1.1 for instance. As a result, the discount factor is no longer needed a priori. This is addressed in Chapter 2. In Chapter 3, we pursue this direction further and provide *running-cost* guarantees, that is how close the actual ensued cost along the solution to the controlled system is related to the infinite-horizon optimal cost (1.13). However, the algorithm stays unchanged: stability is only exploited for the analysis.

Tailored algorithms. This is the main objective of Chapter 3, where we study, *and modify*, an algorithm called Optimistic Planning (OP) [32]. OP is similar to VI, as both virtually optimize finite-horizon discounted costs as in (1.24) to approximate the infinite-horizon one (1.13). OP is specialized to nonlinear systems with a finite number of inputs and is a planning-based approach. This technique is therefore relevant for the near-optimal control of nonlinear switched systems, for which the switching signal is the control. While VI in general incurs approximation errors, OP is able to exactly calculate the optimal cost for a finite-horizon control problem. Nevertheless, OP suffers from similar limitations as those often encountered with VI [7], namely it assumes bounded stage costs and requires a discount factor. These requirements can be relaxed by exploiting stability. This is what we demonstrate in Chapter 3 and we call the modified version of OP, OP_{\min} . Compared to OP, where the near-optimality bounds are limited to γ and horizon d , OP_{\min} exploits stability to its benefit and results in tighter near-optimality bounds. Moreover, OP_{\min} can vary the horizon “on-the-fly”, which helps reducing its computational costs.

1.9 Organization and summary of the contributions

The rest of the thesis is organized as follows.

In Chapter 2, we analyse the stability of nonlinear discrete-time systems whose inputs are generated by an algorithm that minimizes a discounted finite-horizon cost, like in VI. Based on general stabilizability and detectability conditions inspired from [24, 55], we show that semiglobal practical stability properties can be ensured, where the adjustable

parameters are the horizon length and the discount factor. This implies that, for any given ball of initial conditions and any neighborhood of the attractor, by taking both the discount factor γ and the number of iteration d sufficiently large, solutions initialized in this ball converge to the given neighborhood of the attractor. The considered assumptions relaxes those in [30], so that more general systems and stage costs can be considered, and the results allow to investigate general set stability. Stronger stability properties are then derived by strengthening the assumptions. Bounds on the number of iterations, i.e. the horizon length, and the discount factor are provided, which are often less restrictive than those found in [24, 55], when there is no discount factor or when the horizon is infinite, respectively. We also provide new relationships between the optimal value functions of the discounted, undiscounted, infinite-horizon and finite-horizon costs respectively, whose bounds differs from those found in the ADP literature, e.g. Lemma 1.1. Lastly, stability results are also provided for systems controlled by AVI.

In Chapter 3, we consider OP [32]. As mentioned in Section 1.8, this algorithm is relevant for the optimal control of nonlinear switched systems, where the control input is the switching signal. We revisit and modify the original algorithm presented in [32] to be exploitable in control. Namely, we aim at minimizing, and not maximizing, a cost. The latter does not have to be discounted, contrary to [32] where this is essential, and the stage cost does not have to be bounded like in Lemma 1.1, which allows considering quadratic costs for example. Based on the results of Chapter 2, we provide general stability results for systems controlled by the new version of OP we propose, called OP_{\min} . We then exploit stability to demonstrate that the obtained near-optimality bound (hence, OP_{\min}) has major advantages compared to the bound originally given by OP. The results are applied to the situation where we are given a finite number of feedback controllers, among which one is globally stabilizing, and we need to optimally switch among these, similarly to what is done in uniting control e.g., [57–59]. As a result of this chapter, OP_{\min} provides a new tool for the near-optimal, stable control of nonlinear switched discrete-time systems for generic cost functions.

In Chapter 4, we briefly revisit the technical results of Chapters 2 and 3 through the

notion of regret, which we introduce and which refers to the cost consequences of near-optimal, “regrettable”, actions. This notion helps to interpret the relationship between near-optimality, closed-loop stability and running cost guarantees viewed in Chapters 2 and 3

In Chapter 5, we provide the conclusions and future perspectives.

The appendices present extra results related to Chapters 2 and 3, which were sidelined as to not hurt the presentation of this thesis. In Appendix D, we also provide a parallel work on AVI, with a slightly different approach compared to Chapter 2.

Chapter 2. Finite-horizon discounted optimal control

2.1	Introduction	15
2.2	Problem statement	16
2.3	Stability results	20
2.4	Relationships between optimal value functions	35
2.5	Near-optimal sequence of inputs	39
2.6	Illustrative example: an inverted pendulum	47
2.7	Conclusion	50

2.1 Introduction

As explained in the previous chapter, discounted costs are ubiquitous in fields like Approximate Dynamic Programming (ADP) and Reinforcement Learning (RL), which provide a variety of efficient algorithms to construct near-optimal inputs for nonlinear discrete-time systems. The related literature concentrates on optimality and ignores the stability and robustness issue.

In this chapter, we analyse the stability of nonlinear discrete-time systems whose inputs are generated by an algorithm that minimizes a discounted finite-horizon cost, as it appears to be the case for many optimization techniques including Value Iteration (VI), for

instance as explained in Chapter 1.7. For this purpose, assumptions are made related on the stabilizability of the system and its detectability with respect to the stage cost. Then, a Lyapunov function for the closed-loop system is constructed and a uniform semiglobal stability property is ensured, where the adjustable parameters are both the discount factor and the horizon length. Stronger stability properties such as global exponential stability are also provided by strengthening the initial assumptions. We give bounds on the discount factor and the horizon length under which stability holds and we show on examples that these are less conservative than the bounds of the literature for discounted infinite-horizon cost and undiscounted finite-horizon costs, respectively.

In addition, we provide new relationships between the optimal value functions of the discounted, undiscounted, infinite-horizon and finite-horizon costs respectively, which appear to be very different from those available in the approximate dynamic programming literature. These relationships rely on assumptions that are more likely to be satisfied in a control context. Finally, we investigate stability when only a near-optimal sequence of inputs for the discounted finite-horizon cost is available, covering approximate VI as a particular case.

The results of this chapter corresponds to the publications [21, 22].

2.2 Problem statement

Consider the system

$$x^+ = f(x, u), \tag{2.1}$$

with state $x \in \mathbb{R}^n$, input $u \in \mathcal{U}(x) \subseteq \mathbb{R}^m$, where $\mathcal{U}(x)$ is the nonempty set of admissible inputs for state x , and $f : \mathcal{W} \rightarrow \mathbb{R}^n$ where $\mathcal{W} := \{(x, u) : x \in \mathbb{R}^n, u \in \mathcal{U}(x)\}$. We use $\phi(k, x, \mathbf{u}|_k)$ to denote the solution to system (2.1) at time $k \in \mathbb{Z}_{\geq 0}$ with initial condition x and inputs $\mathbf{u}|_k = [u_0, u_1, \dots, u_{k-1}]$, with the convention $\phi(0, x, \cdot) = \phi(0, x, \emptyset) = x$.

We study discounted finite-horizon costs of the form

$$J_{\gamma, d}(x, \mathbf{u}) := \sum_{k=0}^d \gamma^k \ell(\phi(k, x, \mathbf{u}|_k), u_k) \tag{2.2}$$

where $x \in \mathbb{R}^n$, \mathbf{u} is a sequence of $d + 1$ admissible inputs, $\ell : \mathcal{W} \rightarrow \mathbb{R}_{\geq 0}$, $\gamma \in (0, 1]$ is the discount factor and $d \in \mathbb{Z}_{>0} \cup \{\infty\}$ is the horizon.

We assume that for any $x \in \mathbb{R}^n$, $\gamma \in (0, 1]$ and $d \in \mathbb{Z}_{>0} \cup \{\infty\}$, there is a sequence \mathbf{u} that minimizes cost (2.2), as formalized next.

Standing Assumption (SA). For any $x \in \mathbb{R}^n$, $\gamma \in (0, 1]$ and $d \in \mathbb{Z}_{>0} \cup \{\infty\}$, there exists a sequence of $d + 1$ admissible inputs $\mathbf{u}_{\gamma,d}^*(x)$, called *optimal input sequence*, which minimizes (2.2), i.e.

$$J_{\gamma,d}(x, \mathbf{u}_{\gamma,d}^*(x)) = \min_{\mathbf{u}} J_{\gamma,d}(x, \mathbf{u}) =: V_{\gamma,d}(x), \quad (2.3)$$

where $V_{\gamma,d}$ is the *optimal cost function*. □

Conditions to ensure the satisfaction of SA can be found in [39]. According to SA, for any $x \in \mathbb{R}^n$, $\gamma \in (0, 1]$ and $d \in \mathbb{Z}_{>0} \cup \{\infty\}$, the set below is non-empty

$$\begin{aligned} \mathcal{U}_{\gamma,d}^*(x) &:= \{u_0 : \exists u_1, \dots, u_d \in \mathbb{R}^m \text{ admissible such that} \\ &V_{\gamma,d}(x) = J_{\gamma,d}(x, [u_0, \dots, u_d])\}. \end{aligned} \quad (2.4)$$

Note that $\mathcal{U}_{\gamma,d}^*(x)$ may be a set with multiple elements because the optimal sequence may be non-unique for given x , γ and d .

We consider the scenario where system (2.1) is controlled in a receding horizon fashion in the sense that, at each time instant $k \in \mathbb{Z}_{\geq 0}$, the first element of the optimal sequence $\mathbf{u}_{\gamma,d}^*(x)$, which may be non-unique, is applied to system (2.1). This leads to the difference inclusion

$$x^+ \in f(x, \mathcal{U}_{\gamma,d}^*(x)) =: F_{\gamma,d}^*(x), \quad (2.5)$$

where $f(x, \mathcal{U}_{\gamma,d}^*(x))$ is the set $\{f(x, u) : u \in \mathcal{U}_{\gamma,d}^*(x)\}$. Note that $\mathcal{U}_{\gamma,d}^*(x)$ is recalculated at each time-step, hence we have an autonomous system. We denote by $\phi(k, x)$, with some abuse of notation, a solution to (2.5) at time $k \in \mathbb{Z}_{\geq 0}$ with initial condition $x \in \mathbb{R}^n$.

Our main objective is to analyse the stability and robustness of system (2.5) using Lyapunov-based arguments. In particular, we want to investigate the influence of the cost parameters γ and d on stability. To this end, we make the following assumptions, inspired by [24], on the stabilizability and detectability of system (2.1) and cost function (2.2).

Assumption 2.1. There exist $\bar{\alpha}_V, \alpha_W \in \mathcal{K}_\infty$, continuous functions $W, \sigma : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$, $\bar{\alpha}_W : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ continuous, non-decreasing and zero at zero, such that the following conditions hold.

(i) For any $x \in \mathbb{R}^n$, $\gamma \in (0, 1]$ and $d \in \mathbb{Z}_{>0} \cup \{\infty\}$,

$$V_{\gamma,d}(x) \leq \bar{\alpha}_V(\sigma(x)). \quad (2.6)$$

(ii) For any $x \in \mathbb{R}^n$, $u \in \mathcal{U}(x)$,

$$W(x) \leq \bar{\alpha}_W(\sigma(x)) \quad (2.7)$$

$$W(f(x, u)) - W(x) \leq -\alpha_W(\sigma(x)) + \ell(x, u). \quad (2.8)$$

□

Function σ in Assumption 2.1 serves as a measure of the state and will be used to define stability. When investigating the stability of the origin for instance, we typically take $\sigma(x) = |x|$, $\sigma(x) = |x|^2$, or $\sigma(x) = x^T P x$ with P a real symmetric, positive definite matrix, for any $x \in \mathbb{R}^n$. When interested in stability of a non-empty set $\mathcal{A} \subseteq \mathbb{R}^n$, σ can be defined as $\sigma = |\cdot|_{\mathcal{A}}$ for instance, where $|x|_{\mathcal{A}} = \inf\{|z - x| : z \in \mathcal{A}\}$ for any $x \in \mathbb{R}^n$. Item (i) of Assumption 2.1 is related to the asymptotic controllability (stabilizability) of system (2.1) with respect to σ , as noted by Lemma 1 from [24]. In the following, we recall a stronger form of Lemma 1 from [24].

Lemma 2.1. *Suppose there exist constant $M, \lambda > 0$ such that, for any $x \in \mathbb{R}^n$, $\ell(\phi(k, x, \mathbf{u}_\infty(x)|_k), u_k(x)) \leq M e^{-\lambda k} \sigma(x)$ holds for all $k \in \mathbb{Z}_{\geq 0}$ and some infinite length sequence $\mathbf{u}_\infty(x)$. We say that ℓ is global exponentially stabilizable to zero with respect to σ for system (2.1). Then, item (i) of Assumption 2.1 is verified with $\bar{\alpha}_V(\sigma(x)) = \frac{M}{1-e^{-\lambda}} \sigma(x)$.*

□

If Lemma 2.1 is not satisfied, but if $\ell(\phi(k, x, \mathbf{u}_\infty(x)|_k), u_k(x)) \leq \beta(\sigma(x), k)$, we can always modify the stage cost to derive a new optimization problem which verifies Lemma 2.1, see [24, Lemma 1].

Note that we prefer “stabilizability” instead of “asymptotic controlability”. Indeed, stabilizability, e.g. $\ell(\phi(k, x, \mathbf{u}_\infty(x)), u_k(x)) \leq \beta(\sigma(x), k)$, is a stronger property than “controlability”, finding inputs such that $\ell(\phi(k, x, \mathbf{u}_\infty(x)), u_k(x)) \rightarrow 0$. This is akin to how Lyapunov stability is stronger than convergence, see [50]. Moreover, since the upper-bound β is \mathcal{KL} , this in turn means that small variations in $\sigma(x)$ implies bounded variations for $\ell(\phi(k, x, \mathbf{u}_\infty(x)), u_k(x))$, which is also analogous of Lyapunov stability.

On the other hand, item (ii) of Assumption 2.1 is a detectability property of the stage cost ℓ with respect to σ . To see this, consider the particular case where $W = 0$ so that (2.8) reduces to $\alpha_W(\sigma(x)) \leq \ell(x, u)$. Thus, when $\ell(x, u) = 0$, $\sigma(x) = 0$ since $\alpha_W \in \mathcal{K}_\infty$. To illustrate the case $W \neq 0$, we borrow from [33] the following example.

Example 2.1. Consider the following linear system

$$x^+ = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x + \begin{bmatrix} 1/2 \\ 1 \end{bmatrix} u, \quad (2.9)$$

where $x \in \mathbb{R}^2$ and $u \in \mathbb{R}$. Given cost $\ell(x, u) = x^\top \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x + u^2$, item (ii) of SA2 is verified with $\sigma(x) = |x|$, $W(x) = \frac{1}{10}x^\top \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix} x$ and $\alpha_W(\sigma(x)) = \frac{3}{40}(\sigma(x))^2$. ■

Remark 2.1. A more general detectability assumption is made in [24, 55], namely $W(f(x, u)) - W(x) \leq -\alpha_W(\sigma(x)) + \chi(\ell(x, u))$ where $\chi \in \mathcal{K}_\infty$, instead of (2.8). It is possible to obtain stability results in this case, at the price of more technicalities. We have not addressed this case to not compromise the clarity of our main results with the technicalities involved in deriving this more general case. □

Remark 2.2. Throughout the text, we assume $\mathbb{I} - \alpha_W \circ (\bar{\alpha}_V + \bar{\alpha}_W)^{-1} \in \mathcal{K}_\infty$. This is without loss of generality, as, if it is not the case, we can always upper-bound $\mathbb{I} - \alpha_W \circ (\bar{\alpha}_V + \bar{\alpha}_W)^{-1}$ by $\mathbb{I} - \tilde{\alpha}$, which is of class \mathcal{K}_∞ for some suitable $\tilde{\alpha} \in \mathcal{K}_\infty$, according to Lemma B.1 in [37]. This substitution is enough for the forthcoming stability analysis. □

2.3 Stability results

2.3.1 Lyapunov properties

The satisfaction of Assumption 2.1 allows us to derive the following Lyapunov properties, that we use to derive the main stability result for system (2.5) afterwards.

Theorem 2.1. *Suppose Assumption 2.1 holds. There exist $\underline{\alpha}_Y, \bar{\alpha}_Y, \alpha_Y \in \mathcal{K}_\infty$ and, for any $\gamma \in (0, 1]$ and $d \in \mathbb{Z}_{>0} \cup \{\infty\}$, there exists $Y_{\gamma,d} : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$ such that the following hold.*

(i) For any $x \in \mathbb{R}^n$,

$$\underline{\alpha}_Y(\sigma(x)) \leq Y_{\gamma,d}(x) \leq \bar{\alpha}_Y(\sigma(x)). \quad (2.10)$$

(ii) For any $x \in \mathbb{R}^n$, $v \in F_{\gamma,d}^*(x)$,

$$Y_{\gamma,d}(v) - Y_{\gamma,d}(x) \leq \frac{1}{\gamma} \left(-\alpha_Y(\sigma(x)) + \Upsilon(Y_{\gamma,d}(x), \gamma, d) \right) \quad (2.11)$$

where $\Upsilon : \mathbb{R}_{\geq 0} \times (0, 1] \times (\mathbb{Z}_{>0} \cup \{\infty\}) \rightarrow \mathbb{R}_{\geq 0}$ is defined in Table 2.1, and is such that, for any $s \geq 0$, $\Upsilon(s, \gamma, d) \rightarrow 0$ when $\gamma \rightarrow 1$ and $d \rightarrow \infty$. \square

Proof. We distinguish three cases depending on the value of γ and d .

Case 1: $\gamma \in (0, 1)$ and $d \in \mathbb{Z}_{>0}$.

Let $\gamma \in (0, 1)$, $d \in \mathbb{Z}_{>0}$, $x \in \mathbb{R}^n$ and $v \in F_{\gamma,d}^*(x)$. There exists $[u_0^*, u_1^*, \dots, u_d^*] = \mathbf{u}_{\gamma,d}^*(x)$ such that $v = f(x, u_0^*)$ and $\mathbf{u}_{\gamma,d}^*(x)$ is an optimal input sequence for system (2.1) with cost (2.2). Hence $V_{\gamma,d}(x) = J_{\gamma,d}(x, \mathbf{u}_{\gamma,d}^*(x))$.

The proof can be summarized as follows. We start by showing item (ii) of Theorem 2.1. For this, we first upper-bound $V_{\gamma,d}(v)$ using Bellman equation. We then derive a preliminary upper-bound on $V_{\gamma,d}(v) - V_{\gamma,d}(x)$. Afterwards, we define $Y_{\gamma,d} := V_{\gamma,d} + W$ where W comes from Assumption 2.1 and we derive the desired result.

Since stage cost ℓ is nonnegative and in view of item (i) of Assumption 2.1,

$$\ell(x, u_0^*) \leq V_{\gamma,d}(x) \leq \bar{\alpha}_V(\sigma(x)). \quad (2.12)$$

Let $j \in \{1, \dots, d\}$. Consider the sequence $\hat{\mathbf{u}} := [u_1^*, u_2^*, \dots, u_{d-j}^*, \bar{\mathbf{u}}_j]$ where $\bar{\mathbf{u}}_j := \mathbf{u}_{\gamma,j}^*(\phi(d-j+1, x, \mathbf{u}_{\gamma,d}^*(x)|_{d-j+1}))$, $\mathbf{u}_{\gamma,d}^*(x)|_{d-j+1} = [u_0^*, \dots, u_{d-j}^*]$ and ϕ denotes the solution of system (2.1). The sequence $\hat{\mathbf{u}}$ consists of the first $d-j$ elements of $\mathbf{u}_{\gamma,d}^*(x)$

after u_0^* , followed by an optimal input sequence of length $j + 1$ for cost $J_{\gamma,d-j}$ at state $\phi(d - j + 1, x, \mathbf{u}_{\gamma,d}^*(x)|_{d-j+1})$. Note that the sequence $\bar{\mathbf{u}}_j$ exist and minimizes $J_{\gamma,j}(\phi(d - j + 1, x, \mathbf{u}_{\gamma,d}^*(x)|_{d-j+1}), \bar{\mathbf{u}}_j)$ from SA. From the definition of cost $J_{\gamma,d}$ in (2.2) and $V_{\gamma,d}$ in (2.3), $V_{\gamma,d}(v) \leq J_{\gamma,d}(v, \hat{\mathbf{u}}) = J_{\gamma,d-j-1}(v, \hat{\mathbf{u}}|_{d-j}) + \gamma^{d-j} J_{\gamma,j}(\phi(d - j, v, \hat{\mathbf{u}}|_{d-j}), \bar{\mathbf{u}}_j)$. By definition of $\bar{\mathbf{u}}_j$, $V_{\gamma,d}(v) \leq \sum_{k=0}^{d-j-1} \gamma^k \ell(\phi(k, v, \hat{\mathbf{u}}|_k), \hat{\mathbf{u}}_k) + \gamma^{d-j} V_{\gamma,j}(\phi(d - j, v, \hat{\mathbf{u}}|_{d-j}))$. For any $k \in \{0, \dots, d\}$, $\phi(k, v, \hat{\mathbf{u}}|_k) = \phi(k + 1, x, [u_0^*, \hat{\mathbf{u}}]_{k+1}) = \phi(k + 1, x, \mathbf{u}_{\gamma,d}^*(x)|_{k+1})$. Similarly, since $j \in \{1, \dots, d\}$ implies $0 \leq d - j < d$, $\phi(d - j, v, \hat{\mathbf{u}}|_{d-j}) = \phi(d - j + 1, x, \mathbf{u}_{\gamma,d}^*(x)|_{d-j+1})$. Thus

$$\begin{aligned} V_{\gamma,d}(v) &\leq \sum_{k=0}^{d-j-1} \gamma^k \ell(\phi(k + 1, x, \mathbf{u}_{\gamma,d}^*(x)|_{k+1}), [\mathbf{u}_{\gamma,d}^*(x)]_{k+1}) \\ &\quad + \gamma^{d-j} V_{\gamma,j}(\phi(d - j + 1, x, \mathbf{u}_{\gamma,d}^*(x)|_{d-j+1})), \end{aligned} \quad (2.13)$$

where $[\mathbf{u}_{\gamma,d}^*(x)]_{k+1} = u_{k+1}^*$. Using the following shorthand notation, we define the *optimal solution* $\phi_k^* := \phi(k, x, \mathbf{u}_{\gamma,d}^*(x))$ and the *optimal stage cost* $\ell_k^* := \ell(\phi_k^*, u_k^*)$ for $k \in \{0, \dots, d\}$.

Hence,

$$V_{\gamma,d}(v) = V_{\gamma,d}(\phi_1^*) \leq \sum_{k=0}^{d-j-1} \gamma^k \ell_{k+1}^* + \gamma^{d-j} V_{\gamma,j}(\phi_{d-j+1}^*). \quad (2.14)$$

Furthermore from the definition of $V_{\gamma,d}(x)$,

$$V_{\gamma,d}(x) = \sum_{k=0}^d \gamma^k \ell_k^* \quad (2.15)$$

and

$$V_{\gamma,d}(x) - \ell_0^* \geq \gamma \sum_{k=0}^{d-j-1} \gamma^k \ell_{k+1}^*. \quad (2.16)$$

Subtracting (2.15) from (2.14), it follows $V_{\gamma,d}(v) - V_{\gamma,d}(x) \leq -\sum_{k=0}^d \gamma^k \ell_k^* + \sum_{k=0}^{d-j-1} \gamma^k \ell_{k+1}^* + \gamma^{d-j} V_{\gamma,j}(\phi_{d-j+1}^*) = -\ell_0^* + (1 - \gamma) \sum_{k=0}^{d-j-1} \gamma^k \ell_{k+1}^* + \gamma^{d-j} V_{\gamma,j}(\phi_{d-j+1}^*)$. In view of (2.16), we have $\sum_{k=0}^{d-j-1} \gamma^k \ell_{k+1}^* \leq \frac{V_{\gamma,d}(x) - \ell_0^*}{\gamma}$, hence

$$\begin{aligned} V_{\gamma,d}(v) - V_{\gamma,d}(x) &\leq -\ell_0^* - \frac{1 - \gamma}{\gamma} \ell_0^* + \frac{1 - \gamma}{\gamma} V_{\gamma,d}(x) + \gamma^{d-j} V_{\gamma,j}(\phi_{d-j+1}^*) \\ &= -\frac{\ell_0^*}{\gamma} + \frac{1 - \gamma}{\gamma} V_{\gamma,d}(x) + \gamma^{d-j} V_{\gamma,j}(\phi_{d-j+1}^*). \end{aligned} \quad (2.17)$$

From item (i) of Assumption 2.1,

$$\begin{aligned} V_{\gamma,d}(v) - V_{\gamma,d}(x) &\leq -\frac{\ell_0^*}{\gamma} + \frac{1-\gamma}{\gamma}V_{\gamma,d}(x) + \gamma^{d-j}\bar{\alpha}_V(\sigma(\phi_{d-j+1}^*)) \\ &= \frac{1}{\gamma}\left(-\ell_0^* + (1-\gamma)V_{\gamma,d}(x) + \gamma^{d-j+1}\bar{\alpha}_V(\sigma(\phi_{d-j+1}^*))\right). \end{aligned}$$

Adding and subtracting $\frac{1-\gamma}{\gamma}W(x)$,

$$V_{\gamma,d}(v) - V_{\gamma,d}(x) \leq \frac{1}{\gamma}\left(-\ell_0^* - (1-\gamma)W(x) + (1-\gamma)(V_{\gamma,d}(x) + W(x)) + \gamma^{d-j+1}\bar{\alpha}_V(\sigma(\phi_{d-j+1}^*))\right). \quad (2.18)$$

Let $Y_{\gamma,d} = V_{\gamma,d} + W$. In view of item (ii) of Assumption 2.1 and since $\gamma \leq 1$, $\gamma W(v) - W(x) \leq -\alpha_W(\sigma(x)) + \ell_0^*$. Dividing everything by γ and since $\frac{1}{\gamma} = 1 + \frac{1-\gamma}{\gamma}$, $W(v) - W(x) - \frac{1-\gamma}{\gamma}W(x) \leq -\frac{\alpha_W(\sigma(x))}{\gamma} + \frac{\ell_0^*}{\gamma}$. Therefore,

$$W(v) - W(x) \leq \frac{1}{\gamma}\left(-\alpha_W(\sigma(x)) + \ell_0^* + (1-\gamma)W(x)\right). \quad (2.19)$$

In view of (2.18) and (2.19),

$$\begin{aligned} Y_{\gamma,d}(v) - Y_{\gamma,d}(x) &\leq \frac{1}{\gamma}\left(-\ell_0^* - (1-\gamma)W(x) + (1-\gamma)Y_{\gamma,d}(x) \right. \\ &\quad \left. + \gamma^{k^*}\bar{\alpha}_V(\sigma(\phi_{k^*}^*)) - \alpha_W(\sigma(x)) + \ell_0^* + (1-\gamma)W(x)\right) \\ &= \frac{1}{\gamma}\left(-\alpha_W(\sigma(x)) + (1-\gamma)Y_{\gamma,d}(x) + \gamma^{k^*}\bar{\alpha}_V(\sigma(\phi_{k^*}^*))\right) \end{aligned} \quad (2.20)$$

where $k^* := d - j + 1$. A major difficulty compared to [24] is how to bound $\sigma(\phi_{k^*}^*)$ in (2.20) because of the discount factor. For this purpose, we use Theorem A.1 given in the Appendix. From item (ii) of Theorem A.1, it follows that, for any $k \in \{0, \dots, d-1\}$,

$$Y_{\gamma,d-(k+1)}(\phi_{k+1}^*) - Y_{\gamma,d-k}(\phi_k^*) \leq \frac{1}{\gamma}\left(-\alpha_Y(\sigma(\phi_k^*)) + (1-\gamma)Y_{\gamma,d-k}(\phi_k^*)\right).$$

We write

$$Y_{\gamma,d-(k+1)}(\phi_{k+1}^*) \leq \frac{-\alpha_Y \circ \bar{\alpha}_Y^{-1}(Y_{\gamma,d-k}(\phi_k^*)) + Y_{\gamma,d-k}(\phi_k^*)}{\gamma} \quad (2.21)$$

where $\bar{\alpha}_Y = \bar{\alpha}_V + \bar{\alpha}_W$ and $\alpha_Y = \alpha_W$, since $\alpha_Y(\sigma(\phi_k^*)) \geq \alpha_Y \circ \bar{\alpha}_Y^{-1}(Y_{\gamma,d-k}(\phi_k^*))$, which follows from item (i) of Theorem A.1. As explained in Remark 2.2, we can assume $\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1} \in \mathcal{K}_\infty$ without loss of generality. Thus, starting from $Y_{\gamma,d}(x)$ and proceeding

by iteration, we have $Y_{\gamma,d-k^*}(\phi_{k^*}^*) \leq \left(\frac{\mathbb{I}-\alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma}\right)^{(k^*)}(Y_{\gamma,d}(x))$. We apply item (i) of Theorem A.1 and conclude that

$$\sigma(\phi_{k^*}^*) \leq \underline{\alpha}_Y^{-1} \left(\left(\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma} \right)^{(k^*)} (Y_{\gamma,d}(x)) \right). \quad (2.22)$$

It follows from (2.20) and (2.22) that

$$Y_{\gamma,d}(v) - Y_{\gamma,d}(x) \leq \frac{1}{\gamma} \left[-\alpha_W(\sigma(x)) + (1-\gamma)Y_{\gamma,d}(x) + \gamma^{k^*} \bar{\alpha}_V \circ \underline{\alpha}_Y^{-1} \left(\left(\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma} \right)^{(k^*)} (Y_{\gamma,d}(x)) \right) \right].$$

Thus, equation (2.11) is verified with $\alpha_Y = \alpha_W \in \mathcal{K}_\infty$ and $\Upsilon(s, \gamma, k^*) = (1-\gamma)s + \gamma^{k^*} \bar{\alpha}_V \circ \underline{\alpha}_Y^{-1} \left(\left(\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma} \right)^{(k^*)} (s) \right)$. Recall that j is freely selected in $\{1, \dots, d\}$, as a result so is $k^* \in \{1, \dots, d\}$. Note that $\Upsilon(s, \gamma, d) \geq 0$ for any $s \geq 0$, as $0 \leq \mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}$, which follows from $\alpha_Y = \underline{\alpha}_Y \leq \bar{\alpha}_Y$. Let $s \geq 0$, consider $(1-\gamma)s + \bar{\alpha}_V \left(\underline{\alpha}_Y^{-1} \left(\left(\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma} \right)^{(d)} (s) \right) \right) \geq \Upsilon(s, \gamma, d)$. Note that $s - \alpha_Y \circ \bar{\alpha}_Y^{-1}(s) < s$ if $s \neq 0$. Indeed, suppose $s - \alpha_Y \circ \bar{\alpha}_Y^{-1}(s) = s$ and $s \neq 0$, this is only possible if $\alpha_Y \circ \bar{\alpha}_Y^{-1}(s) = 0$, we attain a contradiction. Hence, $s - \alpha_Y \circ \bar{\alpha}_Y^{-1}(s) < s$ when $s > 0$, and zero at zero. Therefore $(1-\gamma)s + \bar{\alpha}_V \left(\underline{\alpha}_Y^{-1} \left(\left(\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma} \right)^{(d)} (s) \right) \right) \rightarrow 0$ when $\gamma \rightarrow 1$ and $d \rightarrow \infty$. Finally, recall that $0 \leq \Upsilon(s, \gamma, d)$. It follows, by the sandwich rule, that $\Upsilon(s, \gamma, d) \rightarrow 0$ when $\gamma \rightarrow 1$ and $d \rightarrow \infty$. Hence, item (ii) of Theorem 2.1 holds.

In view of Assumption 2.1, $Y_{\gamma,d} \leq \bar{\alpha}_Y(\sigma(x))$ with $\bar{\alpha}_Y = \bar{\alpha}_V + \bar{\alpha}_W \in \mathcal{K}_\infty$. From item (ii) of Assumption 2.1, we have $W(x) \geq \alpha_W(\sigma(x)) - \ell(x, u_0^*)$. Associated with (2.12), it follows $Y_{\gamma,d} \geq \alpha_W(\sigma(x)) - \ell(x, u_0^*) + \ell(x, u_0^*) = \alpha_W(\sigma(x))$. Thus $\underline{\alpha}_Y = \alpha_W \in \mathcal{K}_\infty$. Item (i) of Theorem 2.1 is satisfied.

Case 2: $\gamma = 1$ and $d \in \mathbb{Z}_{>0}$

By following the steps of Case 1 with $\gamma = 1$, the desired result is obtained.

Case 3: $\gamma \in (0, 1)$ and $d = \infty$

Let $\gamma \in (0, 1)$, $x \in \mathbb{R}^n$ and $v \in F_{\gamma,\infty}^*(x)$. From Bellman equation, $V_{\gamma,\infty}(x) = \ell_0^* + \gamma V_{\gamma,\infty}(v)$, thus $V_{\gamma,\infty}(v) = \frac{-\ell_0^* + V_{\gamma,\infty}(x)}{\gamma}$. By following the steps of Case 1, the desired result is obtained. ■

Table 2.1: Expressions of the functions used in Theorem 2.1

$Y_{\gamma,d} :=$	$V_{\gamma,d} + W$
$\underline{\alpha}_Y :=$	α_W
$\bar{\alpha}_Y :=$	$\bar{\alpha}_V + \bar{\alpha}_W$
$\alpha_Y :=$	α_W
$\Upsilon(s, \gamma, d) :=$	$\begin{cases} (1 - \gamma)s + \gamma^d \bar{\alpha}_V \circ \underline{\alpha}_Y^{-1} \circ \left(\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma} \right)^{(d)}(s) & \text{when } \gamma \in (0, 1) \text{ and } d \in \mathbb{Z}_{>0} \\ \bar{\alpha}_V \circ \underline{\alpha}_Y^{-1} \circ (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^{(d)}(s) & \text{when } \gamma = 1 \text{ and } d \in \mathbb{Z}_{>0} \\ (1 - \gamma)s & \text{when } \gamma \in (0, 1) \text{ and } d = \infty \end{cases}$

Function $Y_{\gamma,d}$ plays the role of a Lyapunov function in Theorem 2.1, and its expression as well as the expressions of $\alpha_Y, \underline{\alpha}_Y, \bar{\alpha}_Y$ are given in Table 2.1. Item (i) states that it is positive definite and radially unbounded with respect to the set $\{x : \sigma(x) = 0\}$, uniformly in γ and d . Item (ii) of Theorem 2.1 shows that $Y_{\gamma,d}$ strictly decreases along the solutions to (2.5) up to a perturbative term Υ , which can be made as small as desired by selecting γ close to 1 and d big.

We stress that γ has to be selected close to 1 **and** d has to be large in order for Υ to be small in (2.11), which is consistent with previous works on discounted infinite-horizon control [55] and undiscounted finite-horizon control [24] where a similar Lyapunov inequality is given. Theorem 2.1 is actually a generalization of Theorem 1 in [24] to discounted cost and of Theorem 1 in [55] to finite-horizon. The perturbative term Υ differs from the corresponding one in (5) in [24] when $\gamma = 1$, and from the one in item (b) of Theorem 1 in [55] when $d = \infty$, because of the way the Lyapunov analysis is carried out in the proof of Theorem 2.1. The new analysis we propose is motivated by the fact that it leads to different bounds, on d (and γ) under which stability is preserved. This is discussed in more detail in Section 2.3.4.

2.3.2 Main result

We are ready to state the main stability result.

Theorem 2.2. Consider system (2.5) and suppose Assumption 2.1 holds. There exists $\beta \in \mathcal{KL}$ such that for any $\delta, \Delta > 0$, there exist $\gamma^* \in (0, 1)$ and $d^* \in \mathbb{Z}_{>0}$ such that for any $\gamma \in (\gamma^*, 1]$, $d \in (d^*, \infty]$, $x \in \{z \in \mathbb{R}^n : \sigma(z) \leq \Delta\}$, any solution $\phi(\cdot, x)$ to system (2.5) satisfies, for all $k \in \mathbb{Z}_{\geq 0}$

$$\sigma(\phi(k, x)) \leq \max\{\beta(\sigma(x), k), \delta\}. \quad (2.23)$$

□

Proof. Let $\Delta, \delta > 0$, $d \in (d^*, \infty]$ and $\gamma \in (\gamma^*, 1]$, where γ^*, d^* are defined in the following, $x \in \mathbb{R}^n$ be such that $\sigma(x) \leq \Delta$, $v \in F_{\gamma, d}^*(x)$. There exists $\mathbf{u}_{\gamma, d}^*(x)$ with first element u_0^* such that $v = f(x, u_0)$ according to SA.

Define $\tilde{\Delta} := \bar{\alpha}_Y(\Delta)$, $\tilde{\delta} := (\mathbb{I} - \frac{\bar{\alpha}_Y}{2})^{-1} \circ \underline{\alpha}_Y(\delta)$, $\tilde{\alpha}_Y := \alpha_Y \circ \bar{\alpha}_Y^{-1}$, where^{†1} $\alpha_Y, \bar{\alpha}_Y, \underline{\alpha}_Y$ come from Theorem 2.1. Let (γ^*, d^*) such that the following holds for all $d' > d^*, \gamma' \in (\gamma^*, 1]$,

$$\forall s \in [\tilde{\delta}, \tilde{\Delta}], \quad \Upsilon(s, \gamma', d') \leq \left(1 - \frac{\gamma'}{2}\right) \tilde{\alpha}_Y(s), \quad (2.24)$$

where Υ comes from Theorem 2.1. Such a pair (γ^*, d^*) always exists for the following reason. Consider the function $\psi : (\gamma, d) \mapsto \max_{s \in [\tilde{\delta}, \tilde{\Delta}]} (1 - \gamma)s + \bar{\alpha}_V \circ \alpha_W^{-1} \circ \left(\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma}\right)^{(d)}(s)$. Clearly, $\Upsilon(s, \gamma, d) \leq \psi(\gamma, d)$ for any $s \in [\tilde{\delta}, \tilde{\Delta}]$ in view of the definition of Υ , and recall from Theorem 2.1 that $\Upsilon(s, \gamma, d) \geq 0$, thus $0 \leq \Upsilon(s, \gamma, d) \leq \psi(\gamma, d)$. As explained in the proof of Theorem 2.1, $\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1} < \mathbb{I}$ on $\mathbb{R}_{>0}$, hence $\psi(\gamma, d) \rightarrow 0$ as $(\gamma, d) \rightarrow (1, \infty)$. Thus, from the definition of the limit, there exists a pair (γ^*, d^*) in $(0, 1) \times \mathbb{Z}_{>0}$ such that $\psi(\gamma, d) < \frac{1}{2}\tilde{\alpha}_Y(\tilde{\delta})$ for any $(\gamma, d) \in (\gamma^*, 1] \times (d^*, \infty)$. As a result, for any $s \in [\tilde{\delta}, \tilde{\Delta}]$, $\gamma \in (\gamma^*, 1]$ and $d \in (d^*, \infty)$, $\Upsilon(s, \gamma, d) \leq \psi(\gamma, d) < \frac{1}{2}\tilde{\alpha}_Y(\tilde{\delta})$. Meanwhile, the right hand side of (2.24) is bounded below by $\frac{1}{2}\tilde{\alpha}_Y(\tilde{\delta})$, since $\tilde{\alpha}_Y \in \mathcal{K}_\infty$, $s \in [\tilde{\delta}, \tilde{\Delta}]$ and $\gamma \in (0, 1]$, thus $\Upsilon(s, \gamma, d) \leq \frac{1}{2}\tilde{\alpha}_Y(\tilde{\delta}) \leq (1 - \frac{\gamma}{2}) \tilde{\alpha}_Y(s)$ and (2.24) holds.

In view of item (ii) of Theorem 2.1 and since $\alpha_Y(\sigma(x)) \geq \alpha_Y \circ \bar{\alpha}_Y^{-1}(Y_{\gamma, d}(x))$ according to item (i) of Theorem 2.1,

$$Y_{\gamma, d}(v) - Y_{\gamma, d}(x) \leq \frac{1}{\gamma} \left(-\tilde{\alpha}_Y(Y_{\gamma, d}(x)) + \Upsilon(Y_{\gamma, d}(x), \gamma, d) \right). \quad (2.25)$$

^{†1} Note since $\mathbb{I} - \bar{\alpha}_Y = \mathbb{I} - \alpha_Y \circ (\bar{\alpha}_V + \bar{\alpha}_W)^{-1} \in \mathcal{K}_\infty$ as assumed without loss of generality in Remark 2.2. Thus $\mathbb{I} - \frac{\bar{\alpha}_Y}{2} = \mathbb{I} - \bar{\alpha}_Y + \frac{\bar{\alpha}_Y}{2} \in \mathcal{K}_\infty$, hence $(\mathbb{I} - \frac{\bar{\alpha}_Y}{2})^{-1} \in \mathcal{K}_\infty$ and $\tilde{\delta}$ is well defined.

Since $\sigma(x) \leq \Delta$, $Y_{\gamma,d}(x) \leq \bar{\alpha}_Y(\sigma(x)) \leq \bar{\alpha}_Y(\Delta) = \tilde{\Delta}$. Therefore, when $Y_{\gamma,d}(x) \geq \tilde{\delta}$, we derive from (2.24) that $-\tilde{\alpha}_Y(Y_{\gamma,d}(x)) + \Upsilon(Y_{\gamma,d}(x), \gamma, d) \leq -\frac{\gamma}{2}\tilde{\alpha}_Y(Y_{\gamma,d}(x))$. Thus, from (2.25),

$$\begin{aligned} Y_{\gamma,d}(v) - Y_{\gamma,d}(x) &\leq \frac{1}{\gamma} \left(-\tilde{\alpha}_Y(Y_{\gamma,d}(x)) + \Upsilon(Y_{\gamma,d}(x), \gamma, d) \right) \\ &\leq -\frac{1}{2}\tilde{\alpha}_Y(Y_{\gamma,d}(x)). \end{aligned} \quad (2.26)$$

Consider now $Y_{\gamma,d}(x) \in [0, \tilde{\delta})$. From the definition of Υ , note that $\Upsilon(\cdot, \gamma, d) - (1 - \gamma)\mathbb{I} \in \mathcal{K}_\infty$ or zero. It follows then that $\gamma\mathbb{I} - \tilde{\alpha}_Y + \Upsilon(\cdot, \gamma, d) \in \mathcal{K}_\infty$. Indeed, $\gamma\mathbb{I} - \tilde{\alpha}_Y + \Upsilon(\cdot, \gamma, d) = \gamma\mathbb{I} - \tilde{\alpha}_Y + \Upsilon(\cdot, \gamma, d) - (1 - \gamma)\mathbb{I} + (1 - \gamma)\mathbb{I} = \mathbb{I} - \tilde{\alpha}_Y + \Upsilon(\cdot, \gamma, d) - (1 - \gamma)\mathbb{I}$, which is in \mathcal{K}_∞ since $\mathbb{I} - \tilde{\alpha}_Y \in \mathcal{K}_\infty$ as assumed without loss of generality in Remark 2.2 and $\Upsilon(\cdot, \gamma, d) - (1 - \gamma)\mathbb{I} \in \mathcal{K}_\infty$ or zero as noted before. Therefore, for $Y_{\gamma,d}(x) \in [0, \tilde{\delta})$ and in view of (2.25),

$$\begin{aligned} Y_{\gamma,d}(v) &\leq \frac{1}{\gamma} \left(\gamma Y_{\gamma,d}(x) - \tilde{\alpha}_Y(Y_{\gamma,d}(x)) + \Upsilon(Y_{\gamma,d}(x), \gamma, d) \right) \\ Y_{\gamma,d}(v) &\leq \frac{1}{\gamma} \left(\gamma \tilde{\delta} - \tilde{\alpha}_Y(\tilde{\delta}) + \Upsilon(\tilde{\delta}, \gamma, d) \right). \end{aligned}$$

From (2.24), we derive

$$\begin{aligned} Y_{\gamma,d}(v) &\leq \frac{1}{\gamma} \left(\gamma \tilde{\delta} - \tilde{\alpha}_Y(\tilde{\delta}) + \left(1 - \frac{\gamma}{2}\right) \tilde{\alpha}_Y(\tilde{\delta}) \right) \\ Y_{\gamma,d}(v) &\leq \frac{1}{\gamma} \left(\gamma \tilde{\delta} - \frac{\gamma}{2} \tilde{\alpha}_Y(\tilde{\delta}) \right) \\ Y_{\gamma,d}(v) &\leq \tilde{\delta} - \frac{\tilde{\alpha}_Y(\tilde{\delta})}{2}. \end{aligned}$$

Given the definition of $\tilde{\delta}$,

$$Y_{\gamma,d}(v) \leq \left(\mathbb{I} - \frac{\tilde{\alpha}_Y}{2} \right) (\tilde{\delta}) = \underline{\alpha}_Y(\delta). \quad (2.27)$$

Thus, whenever $Y_{\gamma,d}(x) \leq \underline{\alpha}_Y(\delta)$, $Y_{\gamma,d}(v) \leq \underline{\alpha}_Y(\delta)$ follows. Indeed, if $Y_{\gamma,d}(x) \in [\tilde{\delta}, \tilde{\Delta}]$, $Y_{\gamma,d}(v) \leq Y_{\gamma,d}(x) \leq \underline{\alpha}_Y(\delta)$ according to (2.26), and if $Y_{\gamma,d}(x) \in [0, \tilde{\delta})$, we deduce $Y_{\gamma,d}(v) \leq \underline{\alpha}_Y(\delta)$ according to (2.27). Hence the set $\{z \in \mathbb{R}^n : Y_{\gamma,d}(z) \leq \underline{\alpha}_Y(\delta)\}$ is forward invariant^{†2} for system (2.5).

^{†2} The corresponding step was omitted in the proof of Theorem 2 in [55].

We then invoke the same arguments as in the proof of Theorem 2 in [55]. Hence, there exists $\tilde{\beta} \in \mathcal{KL}$ such that for any $k \in \mathbb{Z}_{\geq 0}$ and solution $\phi(k, x)$ to system (2.5),

$$Y_{\gamma,d}(\phi(k, x)) \leq \max\{\tilde{\beta}(Y_{\gamma,d}(x), k), \underline{\alpha}_Y(\delta)\}. \quad (2.28)$$

Finally, using $\underline{\alpha}_Y(\sigma(x)) \leq Y_{\gamma,d}(x) \leq \bar{\alpha}_Y(\sigma(x))$, we obtain

$$\sigma(\phi(k, x)) \leq \max\{\underline{\alpha}_Y^{-1}(\tilde{\beta}(\bar{\alpha}_Y(\sigma(x)), k)), \delta\}. \quad (2.29)$$

Thus (2.23) holds with $\beta(s, k) = \underline{\alpha}_Y^{-1}(\tilde{\beta}(\bar{\alpha}_Y(s), k))$. ■

Theorem 2.2 ensures a semiglobal practical stability property, i.e. given any set of initial conditions of the form $\{z \in \mathbb{R}^n : \sigma(z) \leq \Delta\}$ where $\Delta > 0$, and any (arbitrarily small) δ , we can select γ and d such that (2.23) holds. The key inequality for deriving γ^* and d^* can be found in the proof of Theorem 2.2, see (2.24). A clear relationship between the choice of γ^* and d^* is difficult to draw for Theorem 2.2. Explicit bounds are provided in the sequel by strengthening the conditions of Theorem 2.2, which also allows us to ensure stronger stability properties.

Remark 2.3. Theorem 2.2 holds even when the horizon d of cost (2.2) varies with time as long as it remains larger than d^* as defined in Theorem 2.2. In Chapter 3, we prove and utilize this remark to provide some flexibility on the horizon to be exploited by the optimizing algorithm. □

We know from [40] that it is essential to work with a continuous Lyapunov function to endow the stability properties with some nominal robustness^{†3}. Here, function $Y_{\gamma,d}$ in Theorem 2.2 serves as a Lyapunov function. To ensure it is continuous, we need to guarantee that $V_{\gamma,d}$ is continuous, since $Y_{\gamma,d} = V_{\gamma,d} + W$ and W is continuous according to Assumption 2.1. Additional assumptions are needed for this purpose.

^{†3} As noted in [55], to apply Theorem 2.8 in [40] the set-valued mapping $F_{\gamma,d}^*$ in (2.5) also has to be such that $F_{\gamma,d}^*(x)$ is nonempty and compact for any $x \in \mathbb{R}^n$. Non-emptiness follows from the Standing Assumption. Compactness of $F_{\gamma,d}^*$ proceeds from the compactness of $\mathcal{U}_{\gamma,d}^*(x)$ (when f is continuous, which is assumed to be the case in Lemma 2.2), which is a consequence of the conditions of Lemma 2.2 and the continuity of $V_{\gamma,d}$ proved in this lemma, according to item (a) of Theorem 1.17 in [63].

Assumption 2.2. The following properties hold.

- (i) f and ℓ are continuous.
- (ii) The admissible input set is uniform in x , i.e. $\mathcal{U}(x) = \mathcal{U}$ for all $x \in \mathbb{R}^n$ and some non-empty set $\mathcal{U} \subseteq \mathbb{R}^m$.
- (iii) Either \mathcal{U} is bounded, i.e. there exists a ball of finite radius \mathcal{B} such that $\mathcal{U}(x) \subseteq \mathcal{B}$ for any $x \in \mathbb{R}^n$, or for each compact set \mathcal{C} , $\eta \in \mathbb{R}$, and $d \in \mathbb{Z}_{\geq 0}$, there exists $\mu > 0$ such that for any $x \in \mathcal{C}$, all admissible sequences of inputs \mathbf{u} of length $d + 1$ satisfying $J_{\gamma,d}(x, \mathbf{u}) \leq \eta$ satisfy $|u_k| \leq \mu$ for $k \in \{0, \dots, d\}$. \square

The next lemma ensures the continuity of $V_{\gamma,d}$ when d is finite.

Lemma 2.2. Consider system (2.1) and suppose Assumption 2.2 holds. For any $\gamma \in (0, 1]$ and $d \in \mathbb{Z}_{>0}$, function $V_{\gamma,d}$ is continuous. \square

Proof. The proof consists in showing that the conditions of Theorem 1.17 in [63] are satisfied by $V_{\gamma,d}$. Let $\gamma \in (0, 1]$, and $d \in \mathbb{Z}_{>0}$. Since \mathcal{U} is non-empty for all x and $J_{\gamma,d}$ is a finite sum, $J_{\gamma,d}$ is trivially finite for all x and \mathbf{u} , $J_{\gamma,d}$ is thus a proper function according to the definition in Section 1.A of [63]. Note that $x \mapsto J_{\gamma,d}(x, \mathbf{u})$ is a well defined map since \mathbf{u} is an admissible sequence for any $x \in \mathbb{R}^n$ due to item (ii) of Assumption 2.2. Moreover, $J_{\gamma,d}$ is simply the composition, multiplication and addition of f and ℓ , which are continuous functions from item (i) of Assumption 2.2, thus it follows that $J_{\gamma,d}$ is also a continuous function in x and \mathbf{u} . Since item (iii) of Assumption 2.2 holds, $J_{\gamma,d}$ is level-bounded in \mathbf{u} locally uniformly to x according to Definition 1.16 in [63]. That is, for each $x \in \mathbb{R}^n$ and $a \in \mathbb{R}_{\geq 0}$ there is a neighborhood \mathcal{X} of x such that the set $\{(x, \mathbf{u}) : x \in \mathcal{X}, J_{\gamma,d}(x, \mathbf{u}) \leq a\}$ is bounded in $\mathbb{R}^n \times \mathbb{R}^m$. Finally, SA guarantees the existence of $\mathbf{u}_{\gamma,d}^*(x)$, it follows from item (i) of Assumption 2.2 that with a fixed sequence $\mathbf{u} := \mathbf{u}_{\gamma,d}^*(x)$, $J_{\gamma,d}(\bar{x}, \mathbf{u})$ is continuous in \bar{x} . Altogether, we conclude that $V_{\gamma,d}$ is a continuous function by invoking item (c) of Theorem 1.17 in [63]. \blacksquare

When d is infinite or item (ii) of Assumption 2.2 is not verified, Theorem 3 in [55] provides conditions under which $V_{\gamma,d}$ is continuous.

2.3.3 Stronger stability properties

We first strengthen the conditions of Theorem 2.2 to ensure a uniform semiglobal *asymptotic* stability property.

Corollary 2.1. *Suppose that Assumption 2.1 is satisfied and there exist $L > 0$, $\bar{a}_W \geq 0$, $a_W, \bar{a}_V > 0$ such that $\bar{\alpha}_V(s) \leq \bar{a}_V \cdot s$, $\bar{\alpha}_W(s) \leq \bar{a}_W \cdot s$, $\alpha_W(s) \geq a_W \cdot s$ for any $s \in [0, L]$.*

Let $\Delta > 0$, and select $\gamma^ \in (0, 1]$ and $d^* \in \mathbb{Z}_{>0}$ such that*

$$\left(\frac{1 - \frac{a_W}{\bar{a}_Y}}{\gamma^*} \right)^{d^*} \bar{a}_Y L \leq \alpha_W(L) \quad (2.30)$$

$$1 - \gamma^* + \frac{\bar{a}_V}{a_W} \left(1 - \frac{a_W}{\bar{a}_Y} \right)^{d^*} < \frac{a_W}{\bar{a}_Y}, \quad (2.31)$$

and for any $(\gamma, d) \in (\gamma^, 1] \times (d^*, \infty]$,*

$$\Upsilon(\bar{\alpha}_Y(\Delta), \gamma, d) \leq \left(1 - \frac{\gamma}{2} \right) \tilde{\alpha}_Y(\bar{a}_Y L), \quad (2.32)$$

where $\bar{a}_Y := \bar{a}_V + \bar{a}_W$, $\tilde{\alpha}_Y := \alpha_W \circ (\bar{\alpha}_V + \bar{\alpha}_W)^{-1}$ and Υ is defined in Table 2.1. Then, there exist $\beta \in \mathcal{KL}$ independent of Δ such that, for any $x \in \{z \in \mathbb{R}^n : \sigma(z) \leq \Delta\}$, any solution $\phi(\cdot, x)$ to system (2.5) satisfies, $\sigma(\phi(k, x)) \leq \beta(\sigma(x), k)$ for all $k \in \mathbb{Z}_{\geq 0}$. \square

Proof. Let $\Delta > 0$, $\gamma \in (\gamma^*, 1]$ and $d \in (d^*, \infty]$, $x \in \mathbb{R}^n$ be such that $\sigma(x) \leq \Delta$ and $v \in F_{\gamma, d}^*(x)$. Since Assumption 2.1 holds, we can apply the conclusions of Theorem 2.1. From item (b) of Theorem 2.1, $Y_{\gamma, d}(v) - Y_{\gamma, d}(x) \leq \frac{1}{\gamma} \left(-\alpha_W(\sigma(x)) + \Upsilon(Y_{\gamma, d}(x), \gamma, d) \right)$, where^{†4} $\Upsilon(s, \gamma, d) = (1 - \gamma)s + \gamma^d \bar{\alpha}_V \circ \alpha_W^{-1} \circ \left(\frac{\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1}}{\gamma} \right)^{(d)}(s)$ since $\underline{\alpha}_Y = \alpha_Y = \alpha_W$. We use the following strategy. First, we show that $Y_{\gamma, d}(v) - Y_{\gamma, d}(x) \leq -\epsilon Y_{\gamma, d}(x)$ holds for some $\epsilon > 0$ when $Y_{\gamma, d}(x) \in [0, a_Y L]$ due to (2.31). Then, we show that $Y_{\gamma, d}(v) - Y_{\gamma, d}(x) \leq -\frac{1}{2} \tilde{\alpha}_Y(Y_{\gamma, d}(x))$ holds for $\tilde{\alpha}_Y = \alpha_W \circ \bar{\alpha}_Y^{-1}$ when $Y_{\gamma, d}(x) \in (a_Y L, \bar{\alpha}_Y(\Delta)]$ due to (2.32). To conclude, we combine the two inequalities and we defer to the proof of Theorem 2.2.

We first derive properties based on the sublinear conditions of Corollary 2.1. Since $\bar{\alpha}_V(s) \leq \bar{a}_V \cdot s$ and $\bar{\alpha}_W(s) \leq \bar{a}_W \cdot s$ for $s \in [0, L]$, $\bar{\alpha}_Y(s) = \bar{\alpha}_V(s) + \bar{\alpha}_W(s) \leq (\bar{a}_V + \bar{a}_W) \cdot s = \bar{a}_Y \cdot s$. Thus $s = \bar{\alpha}_Y^{-1}(\bar{\alpha}_Y(s)) \leq \bar{\alpha}_Y^{-1}(\bar{a}_Y \cdot s)$ for $s \in [0, L]$ and, for $s' = \bar{a}_Y \cdot s \in [0, \bar{a}_Y L]$,

$$\frac{s'}{\bar{a}_Y} \leq \bar{\alpha}_Y^{-1}(s'). \quad (2.33)$$

^{†4} The case when $\gamma = 1$ or $d = \infty$ will be discussed later.

Similarly, from $a_W \cdot s \leq \alpha_W(s)$ for $s \in [0, L]$ and taking $s' = \alpha_W(s) \in [0, \alpha_W(L)]$, we derive that

$$\alpha_W^{-1}(s') \leq \frac{s'}{a_W}. \quad (2.34)$$

By composing both sides of (2.33) with α_W , we obtain that $\alpha_W(\frac{s}{\bar{a}_Y}) \leq \alpha_W \circ \bar{\alpha}_Y^{-1}(s)$ for $s \in [0, \bar{a}_Y L]$, and since $\frac{s}{\bar{a}_Y} \in [0, L]$, $a_W \frac{s}{\bar{a}_Y} \leq \alpha_W(\frac{s}{\bar{a}_Y})$, thus

$$\frac{a_W}{\bar{a}_Y} s \leq \alpha_W \circ \bar{\alpha}_Y^{-1}(s). \quad (2.35)$$

Therefore $s - \alpha_W \circ \bar{\alpha}_Y^{-1}(s) \leq \left(1 - \frac{a_W}{\bar{a}_Y}\right) s$, and, for any $s \in [0, \bar{a}_Y L]$,

$$\left(\frac{s - \alpha_W \circ \bar{\alpha}_Y^{-1}(s)}{\gamma}\right) \leq \gamma^{-1} \left(1 - \frac{a_W}{\bar{a}_Y}\right) s. \quad (2.36)$$

On the other hand, according to (2.31), $1 - \gamma^* < \frac{a_W}{\bar{a}_Y}$, thus $1 - \frac{a_W}{\bar{a}_Y} < \gamma^*$. Since $\gamma^* < \gamma$, $1 - \frac{a_W}{\bar{a}_Y} < \gamma$. With this inequality, we derive from (2.36) that $\left(\frac{s - \alpha_W \circ \bar{\alpha}_Y^{-1}(s)}{\gamma}\right) \leq s$ for any $s \in [0, \bar{a}_Y L]$. Thus $\left(\frac{s - \alpha_W \circ \bar{\alpha}_Y^{-1}(s)}{\gamma}\right) \in [0, \bar{a}_Y L]$ for $s \in [0, \bar{a}_Y L]$. We can then apply the first inequality in (2.36) iteratively and obtain $\left(\frac{\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1}}{\gamma}\right)^{(d)}(s) \leq \gamma^{-d} \left(1 - \frac{a_W}{\bar{a}_Y}\right)^d s$ for $s \in [0, \bar{a}_Y L]$. Also from (2.36), $\gamma^* < \gamma$ and $d > d^*$, we have that $\left(\frac{1 - \frac{a_W}{\bar{a}_Y}}{\gamma^*}\right)^{d^*} > \left(\frac{1 - \frac{a_W}{\bar{a}_Y}}{\gamma^*}\right)^d > \gamma^{-d} \left(1 - \frac{a_W}{\bar{a}_Y}\right)^d$. Hence, according to (2.30) and (2.36), for $s \in [0, \bar{a}_Y L]$,

$$\left(\frac{\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1}}{\gamma}\right)^{(d)}(s) \leq \gamma^{-d} \left(1 - \frac{a_W}{\bar{a}_Y}\right)^d s \leq \alpha_W(L). \quad (2.37)$$

Since (2.37) holds, we can invoke (2.34) and obtain for $s \in [0, \bar{a}_Y L]$

$$\alpha_W^{-1} \circ \left(\frac{\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1}}{\gamma}\right)^{(d)}(s) \leq \frac{\gamma^{-d}}{a_W} \left(1 - \frac{a_W}{\bar{a}_Y}\right)^d s. \quad (2.38)$$

Moreover, from (2.37), it follows that, for $s \in [0, \bar{a}_Y L]$,

$$\alpha_W^{-1} \circ \left(\frac{\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1}}{\gamma}\right)^{(d)}(s) \leq \alpha_W^{-1} \circ \alpha_W(L) = L. \quad (2.39)$$

From (2.38) and (2.39), we conclude that $\gamma^d \bar{a}_V \circ \alpha_W^{-1} \circ \left(\frac{\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1}}{\gamma}\right)^{(d)}(s) \leq \gamma^d \bar{a}_V \cdot \alpha_W^{-1} \circ \left(\frac{\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1}}{\gamma}\right)^{(d)}(s) \leq \gamma^d \bar{a}_V \frac{\gamma^{-d}}{a_W} \left(1 - \frac{a_W}{\bar{a}_Y}\right)^d s = \frac{\bar{a}_V}{a_W} \left(1 - \frac{a_W}{\bar{a}_Y}\right)^d s$ for $s \in [0, \bar{a}_Y L]$ holds. Therefore, for $s \in [0, \bar{a}_Y L]$,

$$\Upsilon(s, \gamma, d) \leq (1 - \gamma)s + \frac{\bar{a}_V}{a_W} \left(1 - \frac{a_W}{\bar{a}_Y}\right)^d s. \quad (2.40)$$

Note that (2.40) holds even for $\gamma = 1$ or $d = \infty$, indeed $\Upsilon(s, 1, d) \leq \frac{\bar{a}_V}{a_W} \left(1 - \frac{a_W}{\bar{a}_Y}\right)^d s$ and $\Upsilon(s, \gamma, \infty) = (1 - \gamma)s$ for $s \in [0, \bar{a}_Y L]$.

Since $-\alpha_W(\sigma(x)) \leq -\alpha_W \circ \bar{\alpha}_Y^{-1}(Y_{\gamma,d}(x))$ holds from item (i) of Theorem 2.1, for $Y_{\gamma,d}(x) \in [0, \bar{a}_Y L]$ we derive that $-\alpha_W(\sigma(x)) \leq -\frac{a_W}{\bar{a}_Y} Y_{\gamma,d}(x)$ from (2.35). Applying this inequality and (2.40) to item (ii) of Theorem 2.1, we find $Y_{\gamma,d}(v) - Y_{\gamma,d}(x) \leq \frac{-\frac{a_W}{\bar{a}_Y} + (1-\gamma) + \frac{\bar{a}_V}{a_W} \left(1 - \frac{a_W}{\bar{a}_Y}\right)^d}{\gamma} Y_{\gamma,d}(x)$. For (γ^*, d^*) as defined in (2.31), and since $\gamma \in (\gamma^*, 1]$ and $d > d^*$, it follows

$$0 < 1 - \gamma + \frac{\bar{a}_V}{a_W} \left(1 - \frac{a_W}{\bar{a}_Y}\right)^d < \frac{a_W}{\bar{a}_Y}. \quad (2.41)$$

Consequently, there exist $\epsilon \in \left(0, \frac{a_W}{\bar{a}_Y}\right)$ such that $\left(-\frac{a_W}{\bar{a}_Y} + 1 - \gamma + \frac{\bar{a}_V}{a_W} \left(1 - \frac{a_W}{\bar{a}_Y}\right)^d\right) < -\frac{\epsilon}{\gamma} < -\epsilon$. Finally, we conclude that $Y_{\gamma,d}(v) - Y_{\gamma,d}(x) \leq -\epsilon Y_{\gamma,d}(x)$ for $Y_{\gamma,d}(x) \in [0, \bar{a}_Y L]$ since (2.31) holds.

For $Y_{\gamma,d}(x) \in (\bar{a}_Y L, \bar{\alpha}_Y(\Delta)]$, the existence of γ^* and d^* such that (2.32) holds follows from the same arguments as for the existence of γ^* and d^* such that (2.24) holds in the proof of Theorem 2.2, with $\tilde{\Delta} := \bar{\alpha}_Y(\Delta)$ and $\tilde{\delta} := \bar{a}_Y L$. By following the steps of the proof of Theorem 2.2 for (2.26) we obtain that $Y_{\gamma,d}(v) - Y_{\gamma,d}(x) < -\frac{1}{2}\tilde{\alpha}_Y(Y_{\gamma,d}(x))$ for $Y_{\gamma,d}(x) \in (\bar{a}_Y L, \bar{\alpha}_Y(\Delta)]$.

We have found that $Y_{\gamma,d}(v) - Y_{\gamma,d}(x)$ decreases for all $Y_{\gamma,d}(x) \in (0, \bar{\alpha}_Y(\Delta)]$. In particular, by $-\epsilon Y_{\gamma,d}(x)$ for $Y_{\gamma,d}(x) \in [0, \bar{a}_Y L]$ and by $-\frac{1}{2}\tilde{\alpha}_Y(Y_{\gamma,d}(x))$ for $Y_{\gamma,d}(x) \in (\bar{a}_Y L, \bar{\alpha}_Y(\Delta)]$. We conclude the proof by noting that $Y_{\gamma,d}(v) - Y_{\gamma,d}(x) \leq -\min\{\epsilon\mathbb{I}, \frac{1}{2}\tilde{\alpha}_Y\}(Y_{\gamma,d}(x))$. The desired result is then derived by following the final steps of Theorem 2.2. \blacksquare

Corollary 2.1 ensures a uniform semiglobal asymptotic stability property for set $\{x : \sigma(x) = 0\}$, i.e. given the set of initial conditions $\{z \in \mathbb{R}^n : \sigma(z) \leq \Delta\}$ where Δ is any fixed strictly positive real number, we can select γ and d such that (2.23) holds with $\delta = 0$. Consistently with Theorem 2.2, to find a suitable pair (γ^*, d^*) , we have to take γ^* close to 1 and d^* large so that (2.30)-(2.32) hold.

Compared to Corollary 2 in [24] and Corollary 1 in [55], where similar results are derived for infinite-horizon discounted cost and finite-horizon undiscounted cost respectively, we require inequality (2.30). This is not a limitation of our analysis, but an omission of

[55] and [24]. Indeed, $f(s) \leq \bar{f}s$ and $g(s) \leq \bar{g}s$ for $s \in [0, L]$ where $\bar{f}, \bar{g} > 0$, does not imply that $f(g(s)) \leq \bar{f}\bar{g}s$ for $s \in [0, L]$, as was used previously in Corollary 1 of [55] and Corollary 2 of [24]. For a counter example, consider $f(s) := s^2$ and $g(s) := 2s$ for $L = 1$. We have $\bar{f} = 1, \bar{g} = 2$, yet $f(g(s)) = 4s^2 > 2s = \bar{f}\bar{g}s$ for $s \in (\frac{1}{2}, 1]$ which is a contradiction. Thus, a correct analysis has to verify that every sub-part of the chain of functions is inside a valid range. Fortunately, this does not lead to a loss of generality of Corollary 2.1, only requiring an extra condition on γ^* and d^* , which can always be verified.

We can formulate stronger stability properties, namely uniform *global exponential* stability, when conditions of Corollary 2.1 hold with $L = \infty$.

Corollary 2.2. *Suppose that Assumption 2.1 is satisfied and there exist $\bar{a}_W \geq 0, a_W, \bar{a}_V > 0$ such that $\bar{\alpha}_V(s) \leq \bar{a}_V \cdot s, \bar{\alpha}_W(s) \leq \bar{a}_W \cdot s, \alpha_W(s) \geq a_W \cdot s$ for any $s \geq 0$. Let γ^*, d^* be such that*

$$1 - \gamma^* + \frac{\bar{a}_V}{a_W} \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W} \right)^{d^*} < \frac{a_W}{\bar{a}_V + \bar{a}_W}. \quad (2.42)$$

Then, there exist $K, \lambda > 0$, such that for any $\gamma \in (\gamma^, 1], d \in (d^*, \infty]$, for any $x \in \mathbb{R}^n$, the solution $\phi(\cdot, x)$ to system (2.5) satisfies $\sigma(\phi(k, x)) \leq K\sigma(x)e^{-\lambda k}$ for all $k \in \mathbb{Z}_{\geq 0}$. \square*

Sketch of Proof. Let $x \in \mathbb{R}^n, v \in F_{\gamma, d}^*(x)$ and $\gamma \in (\gamma^*, 1], d \in (d^*, \infty]$. Since Assumption 2.1 holds with Corollary 2.1 conditions with $L = \infty$, we can use the sublinear developments of Corollary 2.1 everywhere, that is, for all $x \in \mathbb{R}^n$. In particular, we have shown that given (2.31), there exists $\epsilon \in \left(0, \frac{a_W}{\bar{a}_Y}\right)$ such that $Y_{\gamma, d}(v) - Y_{\gamma, d}(x) \leq -\epsilon Y_{\gamma, d}(x)$ holds for $Y_{\gamma, d}(x) \in [0, \bar{a}_V L]$. Similarly in Corollary 2.2 case, we derive from (2.42) that $Y_{\gamma, d}(v) - Y_{\gamma, d}(x) \leq -\epsilon Y_{\gamma, d}(x)$ holds for any $Y_{\gamma, d}(x) \in \mathbb{R}_{\geq 0}$. We now proceed with the same argument as the proof of Corollary 2 in [55]. Let $x \in \mathbb{R}^n$ and denote $\phi(k, x)$ be a corresponding solution to (2.5) at time $k \in \mathbb{Z}_{\geq 0}$, it holds that $Y_{\gamma, d}(\phi(k, x)) \leq (1 - \epsilon)^k Y_{\gamma, d}(x)$. Since $Y_{\gamma, d}(x) \geq \alpha_W(\sigma(x)) \geq a_W \sigma(x)$ and $Y_{\gamma, d}(x) \leq \bar{\alpha}_Y(\sigma(x)) \leq \bar{a}_Y \sigma(x)$, we conclude that Corollary 2.2 holds with $K = \frac{\bar{a}_Y}{a_W} = \frac{\bar{a}_V + \bar{a}_W}{a_W}$ and $\lambda = -\ln(1 - \epsilon)$. \blacksquare

Corollary 2.2 ensures a uniform global exponential stability property of $\{x : \sigma(x) = 0\}$. It also provides explicit conditions on the pair (γ^*, d^*) under which stability is guaranteed. Indeed, we either first fix $\gamma^* \in (\bar{\gamma}, 1]$ with $\bar{\gamma} = 1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}$ and then select d^* such that

(2.42) holds, or we first fix $d^* > \bar{d}$ with $\bar{d} = \lfloor \frac{\ln(\bar{a}_V(\bar{a}_V + \bar{a}_W)/a_W^2)}{-\ln(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W})} \rfloor$ and select γ^* such that (2.42) holds. The resulting pair (γ^*, d^*) is a suitable candidate for (2.42) by construction.

2.3.4 Comparison of the conditions on γ and d with existing results

It is difficult to compare the conditions on γ and d in the general case of Theorem 2.2 with those in [24] when $\gamma = 1$ and those in [55] when $d = \infty$. For this reason, in the next lemma, we compare the bounds derived from Corollary 2.2 either when $\gamma = 1$ or $d = \infty$ with those given in Corollary 3 in [24], which we denote^{†5} as $d_{[24]}$, and in Corollary 2 in [55], which we denote as $\gamma_{[55]}$, respectively.

Lemma 2.3. *Under the conditions of Corollary 2.2, the following holds.*

(i) *When $\bar{a}_W < a_W$, $\bar{\gamma} < \gamma_{[55]} := \frac{\bar{a}_V}{\bar{a}_V + a_W}$ where $\bar{\gamma} = 1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}$ is the bound on γ given by (2.42) when $d = \infty$.*

(ii) *$\bar{d} = -\ln d_{[24]} / \ln(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W})$ where^{†6} $d_{[24]} := \frac{\bar{a}_V(\bar{a}_V + \bar{a}_W)}{a_W^2}$ and $\bar{d} = -\frac{\ln(\frac{\bar{a}_V(\bar{a}_V + \bar{a}_W)}{a_W^2})}{\ln(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W})}$ is the bound on d given by (2.42) when $\gamma = 1$. \square*

Proof. Item (ii) of Lemma 2.3 is derived immediately by substitution, since $\bar{d} = \frac{\ln(\bar{a}_V(\bar{a}_V + \bar{a}_W)/a_W^2)}{-\ln(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W})}$ and $d_{[24]} := \frac{\bar{a}_V(\bar{a}_V + \bar{a}_W)}{a_W^2}$. Item (i) of Lemma 2.3 follows since $\frac{\bar{\gamma}}{\gamma_{[55]}} < 1$ implies $\frac{(\bar{a}_V + \bar{a}_W - a_W)(\bar{a}_V + a_W)}{(\bar{a}_V + \bar{a}_W)\bar{a}_V} = \frac{\bar{a}_V^2 + \bar{a}_V \bar{a}_W + a_W \bar{a}_W - a_W^2}{\bar{a}_V^2 + \bar{a}_V \bar{a}_W} < 1$, which in turn implies $a_W \bar{a}_W < a_W^2$. Item (i) of Lemma 2.3 holds since $a_W > 0$ by assumption of Corollary 2.2. \blacksquare

Item (i) of Lemma 2.3 implies that the minimum discount factor $\bar{\gamma}$ given by Corollary 2.2 when $d = \infty$ is strictly smaller than the bound found in Corollary 2 from [55] when $\bar{a}_W < a_W$. Item (ii) of Lemma 2.3 provides a direct relationship between the estimate horizon $d_{[24]}$ of Corollary 3 from [24] and our minimum horizon estimate \bar{d} . It can therefore be used to infer which bound is tighter.

We illustrate the results of Lemma 2.3 via examples taken from [24, 55].

^{†5} We make a change of variable $N - 1 = d$ to align our cost function $J_{\gamma, d}$ with the one in [24].

^{†6} There is a slight abuse of notation, since \bar{d} and $d_{[24]}$ are supposed to be integers.

Example 2.2. Consider the following discrete cubic integrator as in [24]

$$\begin{aligned}x_1^+ &= x_1 + u \\x_2^+ &= x_2 + u^3\end{aligned}\tag{2.43}$$

where $(x_1, x_2) = x \in \mathbb{R}^2$ and $u \in \mathcal{U} = \mathbb{R}$. We take $\sigma(x) = |x_1|^3 + |x_2|$ and $\ell(x, u) = \sigma(x) + |u|^3$. From [24], Assumption 2.1 holds with $\bar{\alpha}_V = 14\mathbb{I}$, $\bar{\alpha}_W = 0$, $\alpha_W = \mathbb{I}$. Invoking Corollary 2.2 with $\gamma = 1$, we find $\bar{d} = \lfloor \frac{0 - \ln 14^2}{\ln 13 - \ln 14} \rfloor = 71$. In [24], when no terminal cost is considered, $d_{[24]} = 196$. Thus our new analysis provides a 63% improvement compared to [24]. ■

Example 2.3. Consider the following discrete nonholonomic integrator as in [24]

$$\begin{aligned}x_1^+ &= x_1 + u_1 \\x_2^+ &= x_2 + u_2 \\x_3^+ &= x_3 + x_1 u_2 - x_2 u_1\end{aligned}\tag{2.44}$$

where $(x_1, x_2, x_3) = x \in \mathbb{R}^3$ and $(u_1, u_2) = u \in \mathcal{U} = \mathbb{R}^2$. We take $\sigma(x) = x_1^2 + x_2^2 + 10|x_3|$ and $\ell(x, u) = \sigma(x)$. From [24], Assumption 2.1 with $\bar{\alpha}_V = \frac{16}{5}\mathbb{I}$, $\bar{\alpha}_W = 0$ and $\alpha_W = \mathbb{I}$. Invoking Corollary 2.2 with $\gamma = 1$, we obtain $\bar{d} = \lfloor \frac{2 * (\log 5 - \log 16)}{\log 11 - \log 16} \rfloor = 6$. In [24], when no terminal cost is considered, $d_{[24]} = 11$. Thus our new analysis provides a 36% improvement compared to [24]. ■

Our analysis shows promising relaxation of the horizon d compared to [24], which is important for computation. Yet our results do not supersede [24] in the case when a terminal cost is used. For example, in [24], the minimum horizon when considering terminal cost is $d_{[24]} = 3$ for both examples, and a subsequent work [69] has found $d_{[24]} = 1$ for system (2.43) and $d_{[24]} = 0$ for system (2.44), albeit with different stage cost ℓ and measure σ . As explained in the introduction, we have a special interest in the case where no terminal cost is used, since this corresponds to the usual application of VI to infinite-horizon problems. Extending Theorem 2.1 to also consider terminal costs is left for future work.

We now compare our minimum discount $\bar{\gamma}$ to the one found in [55], illustrating a case in which item (i) of Lemma 2.3 ensures an improvement.

Example 2.4. Consider system (2.44) with $\sigma(x) = x_1^2 + x_2^2 + 10|x_3|$ and $\ell(x, u) = \sigma(x) + |u|^2$, for $x \in \mathbb{R}^3$ and $u \in \mathbb{R}^2$. This slightly different stage cost is considered in [55] to show that SA holds in this case. Assumption 2.1 is satisfied with $\bar{\alpha}_V = \frac{22}{5}\mathbb{I}$, $\bar{\alpha}_W = 0$ and $\alpha_W = \mathbb{I}$. Note that $\bar{a}_W < a_W$, so item (i) of Lemma 2.3 applies. Invoking Corollary 2.2 with $d = \infty$, we calculate $\bar{\gamma} = 1 - \frac{5}{22} = 0.773$. In^{†7} [55], we find that $\gamma_{[55]} = \frac{22}{27} = 0.815$. As expected from item (i) of Lemma 2.3, our new analysis provides an improvement of about 6% compared to the bound given in [24]. ■

2.4 Relationships between optimal value functions

As explained in the introduction Chapter 1, often the goal is to minimize either a discounted infinite-horizon cost or an undiscounted finite-horizon cost, but the mentioned algorithms minimize a finite-horizon discounted cost instead. In this context, it is natural to ask what is the relationship between the cost we originally aim at minimizing, i.e. the infinite-horizon discounted one, and the one we actually minimize, i.e. the finite-horizon one. The next theorem provides relationships between these two costs, as well as between the discounted finite-horizon cost and the undiscounted finite-horizon cost. These results are obtained by exploiting Assumption 2.1.

Theorem 2.3. Let $\gamma \in (0, 1]$, $d \in \mathbb{Z}_{>0}$, $x \in \mathbb{R}^n$ and suppose that Assumption 2.1 holds. Then

$$\left\{ \begin{array}{l} V_{\gamma,d}(x) \leq V_{\gamma,\infty}(x) \leq V_{\gamma,d}(x) + \gamma^d v_{\gamma,d}(x) \\ V_{\gamma,d}(x) \leq V_{1,d}(x) \leq V_{\gamma,d}(x) + (1 - \gamma) \sum_{k=1}^d v_{\gamma,k}(x), \end{array} \right. \quad \begin{array}{l} (2.45a) \\ (2.45b) \end{array}$$

where $v_{\gamma,k}(x) := \bar{\alpha}_V \circ \underline{\alpha}_Y^{-1} \circ \left(\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma} \right)^{(k)} \circ \bar{\alpha}_Y(\sigma(x))$ for $k \in \{1, \dots, d\}$ and $\alpha_Y, \underline{\alpha}_Y, \bar{\alpha}_Y \in \mathcal{K}_\infty$ come from Theorem 2.1, see Table 2.1. □

Proof. Let $x \in \mathbb{R}^n$, $\gamma \in (0, 1]$, $d \in \mathbb{Z}_{>0}$. According to SA, the input sequence $[u_0^*, \dots, u_{d-1}^*, u_d^*] = \mathbf{u}_{\gamma,d}^*(x)$ exists and $J_{\gamma,d}(x, \mathbf{u}_{\gamma,d}^*(x)) = V_{\gamma,d}(x)$. The lower bounds on

^{†7} The calculated value in [55] is misstated as $\frac{17}{22} = 0.88$.

$V_{\gamma,\infty}(x)$ and $V_{1,d}(x)$ in (2.45a) and (2.45b), respectively, follow from the definitions of $V_{\gamma,\infty}(x)$ and $V_{1,d}(x)$ in (2.3). The other inequalities are proved in the following.

Define $\phi_k^* := \phi(k, x, \mathbf{u}_{\gamma,d}^*(x)|_k)$ a solution to (2.1) at time $k \in \{0, \dots, d\}$, initialized at x , and define $\ell_k^* := \ell(\phi_k^*, u_k^*)$ the corresponding stage cost at time k . To prove the second inequality in equation (2.45a), consider the infinite sequence $\hat{\mathbf{u}} := [u_0^*, \dots, u_{d-1}^*, \mathbf{u}_{\gamma,\infty}^*(\phi_d^*)]$, where $\mathbf{u}_{\gamma,\infty}^*$ exist according to SA. By definition of $V_{\gamma,\infty}(x)$ and $V_{\gamma,d}(x)$, $V_{\gamma,\infty}(x) \leq J_{\gamma,\infty}(x, \hat{\mathbf{u}}) = J_{\gamma,d-1}(x, [u_0^*, \dots, u_{d-1}^*]) + \gamma^d J_{\gamma,\infty}(\phi_d^*, \mathbf{u}_{\gamma,\infty}^*(\phi_d^*)) \leq V_{\gamma,d}(x) + \gamma^d V_{\gamma,\infty}(\phi_d^*)$. According to item (i) of Assumption 2.1, $V_{\gamma,\infty}(\phi_d^*) \leq \bar{\alpha}_V(\sigma(\phi_d^*))$. To bound $\sigma(\phi_d^*)$, we invoke (A.8) in the appendix. Hence, $V_{\gamma,\infty}(\phi_d^*) \leq \bar{\alpha}_V \circ \underline{\alpha}_Y^{-1} \circ \left(\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma} \right)^{(d)}(Y_{\gamma,d}(x))$. By invoking item (i) of Theorem 2.1, $Y_{\gamma,d}(x) \leq \bar{\alpha}_Y(\sigma(x))$ and we obtain the second inequality of (2.45a).

For (2.45b), consider the finite sequence $\mathbf{u}_{\gamma,d}^*(x)$. Note that

$$V_{1,d}(x) \leq J_{1,d}(x, \mathbf{u}_{\gamma,d}^*(x)) = \sum_{k=0}^d \ell_k^*. \quad (2.46)$$

By applying Bellman principle to $V_{\gamma,d}(x)$, as done in the proof of Theorem A.1, it follows that $V_{\gamma,d-k}(\phi_k^*) = \ell_k^* + \gamma V_{\gamma,d-(k+1)}(\phi_{k+1}^*)$ for $k \in \{0, \dots, d-1\}$ and $V_{\gamma,0}(\phi_d^*) = \ell_d^*$. By summation we have $\sum_{k=0}^d V_{\gamma,d-k}(\phi_k^*) = \sum_{k=0}^d \ell_k^* + \gamma \sum_{k=0}^{d-1} V_{\gamma,d-(k+1)}(\phi_{k+1}^*) V_{\gamma,d}(\phi_0^*) + \sum_{k=1}^d V_{\gamma,d-k}(\phi_k^*) = \sum_{k=0}^d \ell_k^* + \gamma \sum_{k=1}^d V_{\gamma,d-k}(\phi_k^*)$. Since $\phi_0^* = x$,

$$V_{\gamma,d}(x) + (1 - \gamma) \sum_{k=1}^d V_{\gamma,d-k}(\phi_k^*) = \sum_{k=0}^d \ell_k^*. \quad (2.47)$$

Therefore, in view of (2.46), $V_{1,d}(x) \leq V_{\gamma,d}(x) + (1 - \gamma) \sum_{k=1}^d V_{\gamma,d-k}(\phi_k^*)$. The proof is completed by noting that $V_{\gamma,d-k}(\phi_k^*) \leq \bar{\alpha}_V(\sigma(\phi_k^*)) \leq v_{\gamma,k}(x) := \bar{\alpha}_V \circ \underline{\alpha}_Y^{-1} \circ \left(\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma} \right)^{(k)} \circ \bar{\alpha}_Y(\sigma(x))$, which holds in view of (A.8). \blacksquare

Theorem 2.3 provides explicit computable bounds on $V_{\gamma,\infty}$ (and $V_{1,d}$) based on $V_{\gamma,d}(x)$, γ , d and $\sigma(x)$, which can be used to evaluate the mismatch induced by the minimization of a finite-horizon discounted cost instead of an infinite-horizon discounted one (and an undiscounted finite-horizon one). We expect that the inequalities in (2.45) become equality in the limit case, i.e. when $d \rightarrow \infty$ in (2.45a) and $\gamma \rightarrow 1$ in (2.45b). This is obviously true for (2.45b). The case of (2.45a) is addressed in the next lemma, which indeed ensures that $V_{\gamma,d}(x)$ tends to $V_{\gamma,\infty}(x)$ when $d \rightarrow \infty$, for γ sufficiently close to 1.

Lemma 2.4. Let $\Delta > 0$, $x \in \{z \in \mathbb{R}^n : \sigma(z) \leq \Delta\}$. Suppose Assumption 2.1 holds and consider $\alpha_Y, \bar{\alpha}_Y, \underline{\alpha}_Y$ from Theorem 2.1. For $\gamma \in (1 - \frac{\alpha_Y(\Delta)}{\bar{\alpha}_Y(\Delta)}, 1]$, $V_{\gamma,d}(x) + \gamma^d v_{\gamma,d}(x) \rightarrow V_{\gamma,\infty}(x)$ when $d \rightarrow \infty$, where $v_{\gamma,d}(x) = \bar{\alpha}_V \circ \underline{\alpha}_Y^{-1} \circ \left(\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma}\right)^{(d)} \circ \bar{\alpha}_Y(\sigma(x))$. \square

Proof. Let $\Delta > 0$, $x \in \{z \in \mathbb{R}^n : \sigma(z) \leq \Delta\}$ and $\tilde{\Delta} = \bar{\alpha}_Y(\Delta)$. When $\gamma = 1$, it follows that $v_{1,d}(x) = \bar{\alpha}_V \circ \underline{\alpha}_Y^{-1} \circ (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^{(d)} \circ \bar{\alpha}_Y(\sigma(x))$, which can be made as small as desired by noting that $(\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^{(d)}$ is decreasing in d , as explained in the proof of Theorem 2.1. Thus $v_{1,d}(x) \rightarrow 0$ when $d \rightarrow \infty$.

When $\gamma \in [1 - \frac{\alpha_Y(\Delta)}{\bar{\alpha}_Y(\Delta)}, 1)$, it follows that $(1 - \gamma)\bar{\alpha}_Y(\Delta) \leq \alpha_Y(\Delta)$, thus $(1 - \gamma)\tilde{\Delta} \leq \alpha_Y \circ \bar{\alpha}_Y^{-1}(\tilde{\Delta})$ by definition of $\tilde{\Delta}$, and hence $\frac{\tilde{\Delta} - \alpha_Y \circ \bar{\alpha}_Y^{-1}(\tilde{\Delta})}{\gamma} \leq \tilde{\Delta}$. We have obtained that $\left(\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma}\right)^{(d)}(\tilde{\Delta}) \leq \tilde{\Delta}$. On the other hand, since $\sigma(x) \leq \Delta$ and $\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma}, \bar{\alpha}_Y \in \mathcal{K}_\infty$ as assumed without loss of generality in Remark 2.2 and Theorem 2.1 respectively, we have $\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma}(\bar{\alpha}_Y(\sigma(x))) \leq \frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma}(\Delta) \leq \tilde{\Delta}$. Finally, we have that $\gamma^d v_{\gamma,d}(x) \leq \gamma^d \bar{\alpha}_V \circ \underline{\alpha}_Y(\tilde{\Delta})$, thus $\gamma^d v_{\gamma,d}(x) \rightarrow 0$ when $d \rightarrow \infty$.

Since $\gamma^d v_{\gamma,d}(x) \rightarrow 0$ for $\gamma \in [1 - \frac{\alpha_Y(\Delta)}{\bar{\alpha}_Y(\Delta)}, 1]$ when $d \rightarrow \infty$, it follows from the sandwich rule and (2.45a) that $V_{\gamma,d}(x) + \gamma^d v_{\gamma,d}(x) \rightarrow V_{\gamma,\infty}(x)$ when $d \rightarrow \infty$. \blacksquare

When comparing (2.45a) with the corresponding error bounds usually found in ADP [7], we observe significant improvement when γ is close to 1. Indeed, in algorithms like VI, it is commonly assumed that $\ell(x, u)$ is bounded, e.g. $\ell(x, u) \in [0, 1]$ for any $(x, u) \in \mathcal{W}$, which is not the case here. This property is used to derive that $V_{\gamma,\infty}(x) \leq V_{\gamma,d}(x) + \sum_{k=d}^{\infty} \gamma^k \leq V_{\gamma,d}(x) + \frac{\gamma^{d+1}}{1-\gamma}$, see Lemma 1.1 and Chapter 1.3. The term $\frac{\gamma^d}{1-\gamma}$, which serves as a near-optimality bound, clearly diverges to infinity when $\gamma \rightarrow 1$. This is not the case with the bound in (2.45a), which is small whenever d is large, or when $\sigma(x)$ is small, even when γ is close to 1.

The work in [25] also provides relationship between the *undiscounted* infinite-horizon optimal cost and the finite-horizon one. The results of this section rely on different assumptions than [25], namely we do not rely on a relaxed dynamic programming property (see Proposition 2.2 and Assumption 4.2 in [25]), but on stabilizability and detectability properties. Furthermore, we address discounted costs contrary to [25].

Finally, simpler relationships between the optimal value functions can be obtained under the conditions of Corollary 2.2, as stated next.

Corollary 2.3. *Let $d \in \mathbb{Z}_{>0}$, $x \in \mathbb{R}^n$ and suppose that the conditions of Corollary 2.2 are satisfied. The following hold.*

(i) For any $\gamma \in (0, 1]$,

$$V_{\gamma,d}(x) \leq V_{\gamma,\infty}(x) \leq V_{\gamma,d}(x) + \hat{v}_d(x) \quad (2.48)$$

$$\text{where } \hat{v}_d(x) := \frac{\bar{a}_V(\bar{a}_V + \bar{a}_W)}{a_W} \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right)^d \sigma(x).$$

(ii) For any $\gamma \in (1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}, 1]$,

$$V_{\gamma,d}(x) \leq V_{1,d}(x) \leq V_{\gamma,d}(x) + S_v(x) \quad (2.49)$$

$$\text{where } S_v(x) := (1 - \gamma) \frac{\bar{a}_V(\bar{a}_V + \bar{a}_W)^2}{a_W} \frac{\left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right)}{a_W - (1 - \gamma)(\bar{a}_V + \bar{a}_W)} \sigma(x).$$

□

Proof. Let $d \in \mathbb{Z}_{>0}$, $\gamma \in (0, 1]$ and $x \in \mathbb{R}^n$. Item (i) of Corollary 2.3 follows immediately by application of (2.45a) and of the conditions of Corollary 2.2. Indeed, since $v_d(x) = \bar{\alpha}_V \circ \underline{\alpha}_Y^{-1} \circ \left(\frac{\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}}{\gamma}\right)^{(d)} \circ \bar{\alpha}_Y(\sigma(x))$, we have $v_d(x) \leq \gamma^{-d} \frac{\bar{a}_V(\bar{a}_V + \bar{a}_W)}{a_W} \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right)^d \sigma(x)$. Thus $V_{\gamma,\infty} \leq V_{\gamma,d}(x) + \gamma^d v_d(x) \leq V_{\gamma,d}(x) + \frac{\bar{a}_V(\bar{a}_V + \bar{a}_W)}{a_W} \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right)^d \sigma(x) = V_{\gamma,d}(x) + \hat{v}_d(x)$, which is the desired result.

Assume now that $\gamma \in (1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}, 1]$. Since $\gamma \in (0, 1]$, Theorem 2.3 holds. Recall the definition of v_k from (2.45b) and notice that $v_k(x) \leq \gamma^{-k} \frac{\bar{a}_V(\bar{a}_V + \bar{a}_W)}{a_W} \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right)^k \sigma(x)$, in view of the definition of v_k and by direct application of the conditions of Corollary 2.2. Thus, $\sum_{k=1}^d v_k(x) \leq \sigma(x) \frac{\bar{a}_V(\bar{a}_V + \bar{a}_W)}{a_W} \sum_{k=1}^d \gamma^{-k} \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right)^k = \sigma(x) \frac{\bar{a}_V(\bar{a}_V + \bar{a}_W)}{a_W} \sum_{k=1}^{\infty} \gamma^{-k} \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right)^k$. Since $1 - \frac{a_W}{\bar{a}_V + \bar{a}_W} < \gamma$, $\sum_{k=1}^{\infty} \gamma^{-k} \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right)^k = \frac{\gamma^{-1} \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right)}{1 - \gamma^{-1} \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right)} = \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right) \frac{\bar{a}_V + \bar{a}_W}{a_W - (1 - \gamma)(\bar{a}_V + \bar{a}_W)}$. Finally, $(1 - \gamma) \sum_{k=1}^d v_k(x) \leq S_v(x)$ where $S_v(x) := (1 - \gamma) \sigma(x) \frac{\bar{a}_V(\bar{a}_V + \bar{a}_W)}{a_W} \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right) \frac{\bar{a}_V + \bar{a}_W}{a_W - (1 - \gamma)(\bar{a}_V + \bar{a}_W)}$. Thus $V_{1,d} \leq V_{\gamma,d}(x) + S_v(x)$ and the proof is completed. ■

The mismatch $\hat{v}_d(x)$ between $V_{\gamma,d}(x)$ and $V_{\gamma,\infty}(x)$ in (2.48) is linear in $\sigma(x)$ and no longer depends on γ as in (2.45a). Furthermore, $\hat{v}_d(x) \rightarrow 0$ as $d \rightarrow \infty$ since $1 - \frac{a_W}{\bar{a}_V + \bar{a}_W} \in (0, 1)$ as

shown in the proof of Corollary 2.1. Similarly, the mismatch $S_v(x)$ between $V_{\gamma,d}(x)$ and $V_{1,d}(x)$ is linear in $\sigma(x)$ and no longer depends on d as in (2.45b). Moreover, $S_v(x) \rightarrow 0$ as $\gamma \rightarrow 1$ for any $x \in \mathbb{R}^n$.

2.5 Near-optimal sequence of inputs

In this section, we explore whether solving approximately optimal cost (2.3) preserves our previous results in some sense. Hence, we study the case the available inputs sequence is only near-optimal for the discounted finite-horizon cost in the following sense.

Assumption 2.3. There exists a continuous function $\eta : [0, +\infty) \times \mathcal{P} \rightarrow \mathbb{R}_{\geq 0}$, where $\mathcal{P} := ((0, 1] \times \mathbb{Z}_{\geq 0}) \cup ((0, 1) \times (\mathbb{Z}_{\geq 0} \cup \{+\infty\}))$ with $\eta(\cdot, \gamma, d) \in \mathcal{K}_{\infty}$ for any $(\gamma, d) \in \mathcal{P}$, such that, for any $x \in \mathbb{R}^n$,

$$V_{\gamma,d}(x) \leq \widehat{V}_{\gamma,d}(x) \leq V_{\gamma,d}(x) + \eta(\epsilon, \gamma, d).$$

□

Assumption 2.3 means that, for any $x \in \mathbb{R}^n$ and $(\gamma, d) \in \mathcal{P}$, we know a near-optimal sequence of admissible inputs $\hat{\mathbf{u}}_{\gamma,d}(x)$ where $\eta(\epsilon, \gamma, d)$ is the near-optimality bound. The constant error η can be controlled by choosing parameter $\epsilon > 0$ depending on (γ, d) as $\eta(\cdot, \gamma, d)$ is of class \mathcal{K}_{∞} . Assumption 2.3 covers near-optimality bounds of the form $\eta(\epsilon, \gamma, d) = \frac{1-\gamma^d}{1-\gamma}\epsilon$ where ϵ is related to approximation errors; such bounds are commonly found in the ADP literature, see [7]. The function η in Assumption 2.3 takes value in $\mathbb{R}_{\geq 0} \times \mathcal{P}$, and not in $\mathbb{R}_{\geq 0} \times [0, 1] \times (\mathbb{Z}_{\geq 0} \cup \{+\infty\})$ as we might expect, to cover this type of bounds, which explodes when $(\gamma, d) \rightarrow (1, \infty)$.

Remark 2.4. Note that the bound $\frac{1-\gamma^d}{1-\gamma}\epsilon$ is not necessarily in disagreement with bounds of form $\frac{\gamma^{d+1}}{1-\gamma}\epsilon$ given in Lemma 1.1 for VI or similarly for AVI in Chapter 1.4. Indeed, the near-optimality bound here is between the approximated finite-horizon discounted cost and the optimal finite-horizon discounted cost, while the ones given in Chapter 1 are for the *infinite*-horizon one. □

Compared to [31], where stability is analysed for control inputs generated by Approximate VI (AVI), we: (i) study stability of a generic closed set $\{x \in \mathbb{R}^n : \sigma(x) = 0\}$, hence not restricted to the origin; (ii) do not need the explicit knowledge of a globally stabilizing policy; (iii) address discounted costs; (iv) the stage cost is not necessarily quadratic; (v) allow for a constant error $\eta(\epsilon, \gamma, d)$, which therefore does not depend on x .

We write system (2.1) in closed-loop with a near-optimal sequence of inputs as

$$x^+ \in f(x, \widehat{\mathcal{U}}_{\gamma, d}(x)) =: \widehat{F}_{\gamma, d}(x), \quad (2.50)$$

and we denote by $\hat{\phi}(k, x)$ a solution to (2.50) at time $k \in \mathbb{Z}_{\geq 0}$ with initial condition $x \in \mathbb{R}^n$.

The next theorem provides conditions under which the stability of system (2.5) follows.

Theorem 2.4. *Consider system (2.50) and suppose Assumptions 2.1 and 2.3 hold. There exists $\beta \in \mathcal{KL}$ such that for any $\delta, \Delta > 0$, there exist $(\gamma^*, d^*) \in (0, 1) \times \mathbb{Z}_{\geq 0}$ such that for any $(\gamma, d) \in \mathcal{P}_{\gamma^*, d^*} := ((\gamma^*, 1] \times (d^*, \infty]) \cap \mathcal{P}$, there exists $\epsilon^* > 0$, such that for any $\epsilon \in [0, \epsilon^*)$ and $x \in \{z \in \mathbb{R}^n : \sigma(z) \leq \Delta\}$, any solution $\hat{\phi}(\cdot, x)$ to system (2.50) satisfies*

$$\sigma(\hat{\phi}(k, x)) \leq \max\{\beta(\sigma(x), k), \delta\} \quad \forall k \in \mathbb{Z}_{\geq 0}. \quad (2.51)$$

Moreover, when $\eta(\epsilon, \gamma, d)$ is non-increasing in γ and d , given $\delta, \Delta > 0$, there exists $(\epsilon^*, \gamma^*, d^*) \in \mathbb{R}_{>0} \times (0, 1) \times \mathbb{Z}_{\geq 0}$ such that for any $(\epsilon, \gamma, d) \in [0, \epsilon^*) \times \mathcal{P}_{\gamma^*, d^*}$, (2.51) holds. \square

Proof. The proof consist in building a Lyapunov function for the near-optimal cost, and to conclude by following similar steps as in the proof of Theorem 2.2, by suitable selection of ϵ . Let $x \in \mathbb{R}^n$, $(\gamma, d) \in \mathcal{P}$ and $\epsilon > 0$. Per Assumption 2.3, there exists a near-optimal input sequence $[\hat{u}_0, \hat{u}_1, \dots, \hat{u}_d] = \widehat{\mathbf{u}}_{\gamma, d}(x)$, such that $\hat{v} = f(x, \hat{u}_0)$ and $\widehat{V}_{\gamma, d}(x) = J_{\gamma, d}(x, \widehat{\mathbf{u}}_{\gamma, d}(x))$. Let $k \in \{0, \dots, d\}$ and define $\hat{\phi}_k := \phi(k, x, \widehat{\mathbf{u}}_{\gamma, d}(x)|_k)$ the solution to (2.1) with input $\widehat{\mathbf{u}}_{\gamma, d}(x)$ initialized at x , $\hat{\ell}_k := \ell(\hat{\phi}_k, \hat{u}_k)$ and $\widehat{V}_{\gamma, d-k}(\hat{\phi}_k) := J_{\gamma, d-k}(\hat{\phi}_k, [\hat{u}_k, \dots, \hat{u}_d])$.

By definition of $\widehat{V}_{\gamma, d}$ and $\hat{\ell}_k$, it follows from (2.2) that $\widehat{V}_{\gamma, d-k}(\hat{\phi}_k) = \hat{\ell}_k + \gamma \widehat{V}_{\gamma, d-(k+1)}(\hat{\phi}_{k+1})$ for $k \in \{0, \dots, d-1\}$ and $\widehat{V}_{\gamma, 0}(\hat{\phi}_d) = \hat{\ell}_d$. Invoking item (ii) of Assumption 2.1, we derive $W(\hat{\phi}_k) \geq -\hat{\ell}_k + \alpha_W(\sigma(\hat{\phi}_k)) + W(\hat{\phi}_{k+1})$. Thus, given $\widehat{Y}_{\gamma, d-k}(\hat{\phi}_k) := \widehat{V}_{\gamma, d-k}(\hat{\phi}_k) + W(\hat{\phi}_k)$, we derive $\widehat{Y}_{\gamma, d-k}(\hat{\phi}_k) \geq \alpha_W(\sigma(\hat{\phi}_k)) + \gamma \widehat{Y}_{\gamma, d-(k+1)}(\hat{\phi}_{k+1}) + (1 - \gamma)W(\hat{\phi}_{k+1})$. For $k = d$, we derive similarly that $\widehat{Y}_{\gamma, 0}(\hat{\phi}_d) \geq \alpha_W(\sigma(\hat{\phi}_d))$, thus for any $k \in \{0, \dots, d\}$, it follows that

$$\widehat{Y}_{\gamma, d-k}(\hat{\phi}_k) \geq \alpha_W(\sigma(\hat{\phi}_k)). \quad (2.52)$$

On the other hand, from $\widehat{Y}_{\gamma,d-k}(\hat{\phi}_k) \geq \alpha_W(\sigma(\hat{\phi}_k)) + \gamma \widehat{Y}_{\gamma,d-(k+1)}(\hat{\phi}_{k+1}) + (1-\gamma)W(\hat{\phi}_{k+1})$ for $k \in \{0, \dots, d-1\}$ established above, we deduce $\widehat{Y}_{\gamma,d-k}(\hat{\phi}_k) \geq \alpha_W(\sigma(\hat{\phi}_k)) + \gamma \widehat{Y}_{\gamma,d-(k+1)}(\hat{\phi}_{k+1})$ since $W \geq 0$, which implies

$$\widehat{Y}_{\gamma,d-(k+1)}(\hat{\phi}_{k+1}) \leq \frac{\widehat{Y}_{\gamma,d-k}(\hat{\phi}_k) - \alpha_W(\sigma(\hat{\phi}_k))}{\gamma}. \quad (2.53)$$

We now derive an upper-bound for $\widehat{Y}_{\gamma,d-k}(\hat{\phi}_k)$ in terms of $\widehat{Y}_{\gamma,d}(x)$. Note that $\widehat{Y}_{\gamma,d-(k+1)}(\hat{\phi}_{k+1}) \leq \frac{\widehat{Y}_{\gamma,d-k}(\hat{\phi}_k)}{\gamma}$ holds. In view of Assumptions 2.1 and 2.3, we have that $\widehat{V}_{\gamma,d-k}(\hat{\phi}_k) + W(\hat{\phi}_k) \leq \bar{\alpha}_V(\sigma(\hat{\phi}_k)) + \eta(\epsilon, \gamma, d) + \bar{\alpha}_W(\sigma(\hat{\phi}_k))$, that is, $\widehat{Y}_{\gamma,d-k}(\hat{\phi}_k) \leq \eta(\epsilon, \gamma, d) + \bar{\alpha}_Y(\sigma(\hat{\phi}_k))$, where $\bar{\alpha}_Y := \bar{\alpha}_V + \bar{\alpha}_W$. When $\bar{\alpha}_Y(\sigma(\hat{\phi}_k)) \geq \eta(\epsilon, \gamma, d)$, it follows that $\bar{\alpha}_Y^{-1}(\frac{\widehat{Y}_{\gamma,d-k}(\hat{\phi}_k)}{2}) \leq \sigma(\hat{\phi}_k)$, which implies

$$\tilde{\alpha}_Y(\widehat{Y}_{\gamma,d-k}(\hat{\phi}_k)) \leq \alpha_W(\sigma(\hat{\phi}_k)), \quad (2.54)$$

where $\tilde{\alpha}_Y(s) := \alpha_W \circ \bar{\alpha}_Y^{-1}(\frac{s}{2})$ for all $s \geq 0$. When $\bar{\alpha}_Y(\sigma(\hat{\phi}_k)) \leq \eta(\epsilon, \gamma, d)$, it follows that

$$\widehat{Y}_{\gamma,d-k}(\hat{\phi}_k) \leq 2\eta(\epsilon, \gamma, d). \quad (2.55)$$

From (2.53), (2.54) and (2.55), we derive $\widehat{Y}_{\gamma,d-(k+1)}(\hat{\phi}_{k+1}) \leq \max \left\{ \left(\frac{\mathbb{I} - \tilde{\alpha}_Y}{\gamma} \right) (\widehat{Y}_{\gamma,d-k}(\hat{\phi}_k)), 2\frac{\eta(\epsilon, \gamma, d)}{\gamma} \right\}$. For the sake of convenience, we introduce parameter $\mu > 0$ such that such that $2\frac{\eta(\epsilon, \gamma, d)}{\gamma} \leq \mu$ and that $1 - \frac{\tilde{\alpha}_Y(\mu)}{\mu} \leq \gamma$. Parameter μ can be introduced without loss of generality, since the first condition $2\frac{\eta(\epsilon, \gamma, d)}{\gamma} \leq \mu$ can be always verified by taking ϵ small, and the second is verified by taking γ close to 1. Moreover, we will require later in the proof analogous conditions on γ and η to guarantee stability. Thus,

$$\widehat{Y}_{\gamma,d-(k+1)}(\hat{\phi}_{k+1}) \leq \max \left\{ \left(\frac{\mathbb{I} - \tilde{\alpha}_Y}{\gamma} \right) (\widehat{Y}_{\gamma,d-k}(\hat{\phi}_k)), \mu \right\}. \quad (2.56)$$

We now show that $\frac{\mathbb{I} - \tilde{\alpha}_Y}{\gamma} \in \mathcal{K}_\infty$. Indeed, as explained in Remark 2.2, we can assume without loss of generality that $\mathbb{I} - \alpha_W \circ (\bar{\alpha}_V + \bar{\alpha}_W) \in \mathcal{K}_\infty$. Thus $2\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1} \in \mathcal{K}_\infty$, therefore $\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1} \circ \frac{\mathbb{I}}{2} = \mathbb{I} - \tilde{\alpha}_Y \in \mathcal{K}_\infty$ and we conclude $\frac{\mathbb{I} - \tilde{\alpha}_Y}{\gamma} \in \mathcal{K}_\infty$. Hence, with one iteration of (2.56) to itself, we obtain $\widehat{Y}_{\gamma,d-(k+2)}(\hat{\phi}_{k+2}) \leq \max \left\{ \left(\frac{\mathbb{I} - \tilde{\alpha}_Y}{\gamma} \right)^{(2)} (\widehat{Y}_{\gamma,d-k}(\hat{\phi}_k)), \left(\frac{\mathbb{I} - \tilde{\alpha}_Y}{\gamma} \right) (\mu), \mu \right\}$. By successive iterations of (2.56) and noting that $\left(\frac{\mathbb{I} - \tilde{\alpha}_Y}{\gamma} \right) (\mu) \leq \mu$ because $1 - \frac{\tilde{\alpha}_Y(\mu)}{\mu} \leq \gamma$, we obtain

$$\widehat{Y}_{\gamma,0}(\hat{\phi}_d) \leq \max \left\{ \left(\frac{\mathbb{I} - \tilde{\alpha}_Y}{\gamma} \right)^{(d)} (\widehat{Y}_{\gamma,d}(x)), \mu \right\}. \quad (2.57)$$

Since (2.52) holds,

$$\sigma(\hat{\phi}_d) \leq \sigma_d(\hat{Y}_{\gamma,d}(x)), \quad (2.58)$$

where $\sigma_d(\hat{Y}_{\gamma,d}(x)) := \max \left\{ \alpha_W^{-1} \circ \left(\frac{1 - \bar{\alpha}_Y}{\gamma} \right)^{(d)} (\hat{Y}_{\gamma,d}(x)), \alpha_W^{-1}(\mu) \right\}$. Note that for $s \in [0, \mu]$,

$$\sigma_d(s) = \alpha_W^{-1}(\mu). \quad (2.59)$$

Consider now cost $\hat{V}_{\gamma,d}(\hat{v}) = \hat{V}_{\gamma,d}(f(x, \hat{u}_0))$, where \hat{u}_0 is the first input of the near-optimal sequence $\hat{\mathbf{u}}_{\gamma,d}(x)$. From Assumption 2.3, $\hat{V}_{\gamma,d}(\hat{v}) \leq V_{\gamma,d}(\hat{v}) + \eta(\epsilon, \gamma, d)$. Let $\mathbf{u} := [\hat{u}_1, \hat{u}_2, \dots, \hat{u}_{d-1}, \bar{\mathbf{u}}_1]$ where $\bar{\mathbf{u}}_1 := \mathbf{u}_{\gamma,1}^*(\hat{\phi}_d)$ is the optimal sequence of inputs of length 2 for state $\hat{\phi}_d$ with cost $V_{\gamma,1}(\hat{\phi}_d)$ and \hat{u}_i are the inputs of the near-optimal sequence. It follows from the optimality of $V_{\gamma,d}(\hat{v})$ that $V_{\gamma,d}(\hat{v}) \leq J_{\gamma,d}(\hat{v}, \mathbf{u}) \leq J_{\gamma,d-2}(\hat{v}, [\hat{u}_1, \dots, \hat{u}_{d-1}]) + \gamma^{d-1}V_{\gamma,1}(\hat{\phi}_d)$. Therefore, according to Assumption 2.3,

$$\hat{V}_{\gamma,d}(\hat{v}) \leq J_{\gamma,d-2}(\hat{v}, [\hat{u}_1, \dots, \hat{u}_{d-1}]) + \gamma^{d-1}V_{\gamma,1}(\hat{\phi}_d) + \eta(\epsilon, \gamma, d). \quad (2.60)$$

On the other hand, we have $\hat{V}_{\gamma,d}(x) = J_{\gamma,d}(x, [\hat{u}_0, \hat{u}_1, \dots, \hat{u}_d]) = \hat{\ell}_0 + \gamma J_{\gamma,d-1}(\hat{v}, [\hat{u}_1, \dots, \hat{u}_d])$, which implies $J_{\gamma,d-2}(\hat{v}, [\hat{u}_1, \dots, \hat{u}_{d-1}]) \leq J_{\gamma,d-1}(\hat{v}, [\hat{u}_1, \dots, \hat{u}_d]) = \frac{\hat{V}_{\gamma,d}(x) - \hat{\ell}_0}{\gamma}$. In view of (2.60),

$$\begin{aligned} \hat{V}_{\gamma,d}(\hat{v}) - \hat{V}_{\gamma,d}(x) &\leq J_{\gamma,d-2}(\hat{v}, [\hat{u}_1, \dots, \hat{u}_{d-1}]) + \gamma^{d-1}V_{\gamma,1}(\hat{\phi}_d) + \eta(\epsilon, \gamma, d) - \hat{V}_{\gamma,d}(x) \\ &\leq \frac{\hat{V}_{\gamma,d}(x) - \hat{\ell}_0}{\gamma} - \hat{V}_{\gamma,d}(x) + \gamma^{d-1}V_{\gamma,1}(\hat{\phi}_d) + \eta(\epsilon, \gamma, d) \\ &\leq \frac{-\hat{\ell}_0 + (1 - \gamma)\hat{V}_{\gamma,d}(x) + \gamma^d V_{\gamma,1}(\hat{\phi}_d) + \gamma\eta(\epsilon, \gamma, d)}{\gamma}. \end{aligned} \quad (2.61)$$

From item (ii) of Assumption 2.1 and following the steps in the proof of Theorem 2.1 to obtain (2.19), we have

$$W(\hat{v}) - W(x) \leq \frac{1}{\gamma} \left(-\alpha_W(\sigma(x)) + \hat{\ell}_0 + (1 - \gamma)W(x) \right). \quad (2.62)$$

Summing equations (2.61) and (2.62), we have

$$\hat{Y}_{\gamma,d}(\hat{v}) - \hat{Y}_{\gamma,d}(x) \leq \frac{-\alpha_W(\sigma(x)) + (1 - \gamma)\hat{Y}_{\gamma,d}(x) + \gamma^d V_{\gamma,1}(\hat{\phi}_d) + \gamma\eta(\epsilon, \gamma, d)}{\gamma}. \quad (2.63)$$

Recall that $\hat{Y}_{\gamma,d}(x) \leq \bar{\alpha}_Y(\sigma(x)) + \eta(\epsilon, \gamma, d)$. Again, we distinguish two cases. First, consider $\bar{\alpha}_Y(\sigma(x)) \geq \eta(\epsilon, \gamma, d)$. From (2.54), $\bar{\alpha}_Y(\hat{Y}_{\gamma,d}(x)) = \alpha_W \circ \bar{\alpha}_Y^{-1}(\frac{\hat{Y}_{\gamma,d}(x)}{2}) \leq \alpha_W(\sigma(x))$, thus

$-\alpha_W(\sigma(x)) \leq -\tilde{\alpha}_Y(\widehat{Y}_{\gamma,d}(x))$. On the other hand, from (i) from Assumption 2.1 and (2.58), $V_{\gamma,1}(\hat{\phi}_d) \leq \bar{\alpha}_V(\sigma(\hat{\phi}_d)) \leq \bar{\alpha}_V \circ \sigma_d(\widehat{Y}_{\gamma,d}(x))$. Moreover, $\gamma\eta(\epsilon, \gamma, d) \leq \frac{\gamma^2}{2}\mu$ since $\eta(\epsilon, \gamma, d) \leq \frac{\gamma}{2}\mu$. Altogether, in view of (2.63),

$$\widehat{Y}_{\gamma,d}(\hat{v}) - \widehat{Y}_{\gamma,d}(x) \leq \frac{-\tilde{\alpha}_Y + \widehat{\Upsilon}(\cdot, \mu, \gamma, d)}{\gamma} \left(\widehat{Y}_{\gamma,d}(x) \right) \quad (2.64)$$

holds with $\widehat{\Upsilon}(s, \mu, \gamma, d) := ((1 - \gamma)\mathbb{I} + \gamma^d \bar{\alpha}_V \circ \sigma_d)(s) + \frac{\gamma^2}{2}\mu$. Second, when $\bar{\alpha}_Y(\sigma(x)) \leq \eta(\epsilon, \gamma, d)$, it follows that $\widehat{Y}_{\gamma,d}(x) \leq 2\eta(\epsilon, \gamma, d)$ in view of $\widehat{Y}_{\gamma,d}(x) \leq \bar{\alpha}_Y(\sigma(x)) + \eta(\epsilon, \gamma, d)$. From item (i) of Assumption 2.1 and (2.58), $V_{\gamma,1}(\sigma(\hat{\phi}_d)) \leq \bar{\alpha}_V \circ \sigma_d(\widehat{Y}_{\gamma,d}(x))$. Since $\widehat{Y}_{\gamma,d}(x) \leq 2\eta(\epsilon, \gamma, d) \leq \mu$, it follows from (2.59) that $V_{\gamma,1}(\sigma(\hat{\phi}_d)) \leq \bar{\alpha}_V \circ \alpha_W^{-1}(\mu)$. In view of (2.63) and $-\alpha_W(\sigma(x)) \leq 0$, we derive $\widehat{Y}_{\gamma,d}(\hat{v}) \leq 2\eta(\epsilon, \gamma, d) + \frac{(1-\gamma)2\eta(\epsilon, \gamma, d) + \gamma^d \bar{\alpha}_V \circ \alpha_W^{-1}(\mu) + \gamma\eta(\epsilon, \gamma, d)}{\gamma}$, that is $\widehat{Y}_{\gamma,d}(\hat{v}) \leq (1 + \frac{\gamma}{2})2\frac{\eta(\epsilon, \gamma, d)}{\gamma} + \gamma^{d-1}\bar{\alpha}_V \circ \alpha_W^{-1}(\mu)$. Therefore,

$$\widehat{Y}_{\gamma,d}(\hat{v}) \leq \nu(\mu, \gamma, d) \quad (2.65)$$

holds with $\nu(\mu, \gamma, d) := (1 + \frac{\gamma}{2})\mu + \gamma^{d-1}\bar{\alpha}_V \circ \alpha_W^{-1}(\mu)$. Note that $\nu(\cdot, \gamma, d)$ is class \mathcal{K}_∞ , and both $\nu(\mu, \cdot, d)$ and $\nu(\mu, \gamma, \cdot)$ can be made as small as desired by reducing μ .

Let $\delta, \Delta > 0$ and^{†8} $\tilde{\Delta} := 2\bar{\alpha}_Y(\Delta)$, $\tilde{\delta} := (\mathbb{I} - \frac{1}{2}\tilde{\alpha}_Y)^{-1} \circ \alpha_W(\delta)$. There exists $(\gamma^*, d^*) \in (0, 1) \times \mathbb{Z}_{\geq 0}$ such that for any $(\gamma, d) \in \mathcal{P}_{\gamma^*, d^*}$, there exists $\epsilon^* > 0$ such that for any $\epsilon \in [0, \epsilon^*)$ the following holds

$$\forall s \in [\tilde{\delta}, \tilde{\Delta}], \widehat{\Upsilon}(s, \mu, \gamma, d) \leq \left(1 - \frac{\gamma}{2}\right) \tilde{\alpha}_Y(s) \quad (2.66)$$

$$\nu(\mu, \gamma, d) \leq (\mathbb{I} - \frac{1}{2}\tilde{\alpha}_Y)(\tilde{\delta}) \quad (2.67)$$

$$1 - \frac{\tilde{\alpha}_Y(\mu)}{\mu} \leq \gamma \quad (2.68)$$

$$\frac{2\eta(\epsilon, \gamma, d)}{\gamma} \leq \mu, \quad (2.69)$$

with $\nu(\mu, \gamma, d) = \frac{1+\frac{\gamma}{2}}{\gamma}\mu + \gamma^{d-1}\bar{\alpha}_V \circ \alpha_W^{-1}(\mu)$, $\widehat{\Upsilon}(s, \mu, \gamma, d) = ((1 - \gamma)\mathbb{I} + \gamma^d \bar{\alpha}_V \circ \sigma_d)(s) + \frac{\gamma^2}{2}\mu$, and $\sigma_d(s) = \max \left\{ \alpha_W^{-1} \circ \left(\frac{\mathbb{I} - \tilde{\alpha}_Y}{\gamma} \right)^{(d)}(s), \alpha_W^{-1}(\mu) \right\}$. The satisfaction of (2.66)-(2.69) is not so easy to see, we therefore prove it step by step. First, note that the right hand side of (2.66) is bounded below by $\frac{1}{2}\tilde{\alpha}_Y(\tilde{\delta})$. On the left side, we have $\widehat{\Upsilon}(s, \mu, \gamma, d) \leq (1-\gamma)\tilde{\Delta} + \bar{\alpha}_V \circ \sigma_d(\tilde{\Delta}) + \frac{\mu}{2} =: \psi(\mu, \gamma, d)$. By the same arguments given in the proof of Theorem

^{†8} As seen in the proof of Theorem 2.2, $\mathbb{I} - \tilde{\alpha}_Y \in \mathcal{K}_\infty$ implies $\mathbb{I} - \frac{\tilde{\alpha}_Y}{2} \in \mathcal{K}_\infty$, thus $\tilde{\delta}$ is well defined.

2.2, we have that $\psi(\mu, \gamma, d) \rightarrow \bar{\alpha}_V \circ \alpha_W^{-1}(\mu) + \frac{\mu}{2}$ as $(\gamma, d) \rightarrow (1, \infty)$, which in turn can be made as small as desired by taking μ small. Thus, like in the proof of Theorem 2.2, there exist a triple $(\mu^*_{(2.66)}, \gamma^*_{(2.66)}, d^*_{(2.66)})$ such that for any $(\mu, \gamma, d) \in [0, \mu^*_{(2.66)}] \times \mathcal{P}_{\gamma^*_{(2.66)}, d^*_{(2.66)}}$, $\widehat{\Upsilon}(s, \mu, \gamma, d) \leq \psi(\mu, \gamma, d) \leq \frac{1}{2}\widetilde{\alpha}_Y(\widetilde{\delta})$ and thus (2.66) holds. Regarding (2.67), note that its right hand side is positive since $\widetilde{\delta} > 0$ and^{†8} $\mathbb{I} - \frac{\widetilde{\alpha}_Y}{2} \in \mathcal{K}_\infty$. We have that $\nu(\mu, \gamma, d) \leq \frac{1+\frac{\gamma}{2}}{\gamma}\mu + \bar{\alpha}_V \circ \alpha_W^{-1}(\mu) =: \psi'(\gamma, \mu)$. Since $\frac{1+\frac{\gamma}{2}}{\gamma}$ is decreasing in γ , we have that ψ' is decreasing in γ and can be made as small as desired by taking μ small. Thus in the same vein as for ψ for (2.66), there exists a triple $(\mu^*_{(2.67)}, \gamma^*_{(2.67)}, d^*_{(2.67)})$ such that for any $(\mu, \gamma, d) \in [0, \mu^*_{(2.67)}] \times \mathcal{P}_{\gamma^*_{(2.67)}, d^*_{(2.67)}}$, $\nu(\mu, \gamma, d) \leq \phi'(\mu, \gamma) \leq (\mathbb{I} - \frac{1}{2}\widetilde{\alpha}_Y)(\widetilde{\delta})$ and thus (2.67) holds. For (2.68), note that $1 - \frac{\widetilde{\alpha}_Y(\mu)}{\mu} < 1$ for $\mu > 0$. Thus, for each $\mu > 0$, there exist $\gamma^*_{(2.68)} \in (0, 1)$ such that for any $\gamma \in (\gamma^*_{(2.68)}, 1]$, (2.68) holds. For (2.69), since $\eta(\cdot, \gamma, d) \in \mathcal{K}_\infty$, $\eta(\epsilon, \gamma, d)$ can be made as small as desired by reducing ϵ in function of (μ, γ, d) . Furthermore, for any (μ, γ, d) , there exists $\epsilon^*_{(2.69)}$ such that for any $\epsilon \in [0, \epsilon^*_{(2.69)})$, $\frac{2\eta(\epsilon, \gamma, d)}{\gamma} \leq \mu$. In sum, we select $(\mu^*, \gamma^*, d^*) \in (0, \min\{\mu_{(2.66)}, \mu_{(2.67)}\}) \times \mathcal{P}_{\max\{\gamma^*_{(2.66)}, \gamma^*_{(2.67)}, \gamma^*_{(2.68)}\}, \max\{d^*_{(2.66)}, d^*_{(2.67)}\}}$ and for each $(\gamma, d) \in \mathcal{P}_{\gamma^*, d^*}$, we take $\epsilon^* \in (0, \epsilon^*_{(2.69)})$. In that way (2.66)-(2.68) are satisfied. This choice of (γ^*, d^*) is able to verify the properties mentioned beforehand. That is, inequalities (2.66)-(2.68) hold for μ^* and for every $\epsilon \leq \epsilon^*_{(2.69)}$ chosen in function of $(\gamma, d) \in \mathcal{P}_{\gamma^*, d^*}$, so is (2.69).

We now follow similar lines as in the proof of Theorem 2.2 to prove (2.51). When $\bar{\alpha}_Y(\sigma(x)) \geq \eta(\epsilon, \gamma, d)$, (2.64) holds. In this case, in view of (2.66), $\widehat{Y}_{\gamma, d}(\hat{v}) - \widehat{Y}_{\gamma, d}(x) \leq -\frac{1}{2}\widetilde{\alpha}_Y(\widehat{Y}_{\gamma, d}(x))$ if $\widehat{Y}_{\gamma, d} \in [\widetilde{\delta}, \widetilde{\Delta}]$, and if $\widehat{Y}_{\gamma, d}(x) \in [0, \widetilde{\delta}]$, we derive^{†9} $\widehat{Y}_{\gamma, d}(\hat{v}) \leq \widetilde{\delta} - \frac{1}{2}\widetilde{\alpha}_Y(\widetilde{\delta}) \leq (\mathbb{I} - \frac{1}{2}\widetilde{\alpha}_Y)(\widetilde{\delta})$. For $\bar{\alpha}_Y(\sigma(x)) \leq \eta(\epsilon, \gamma, d)$, (2.65) holds. Hence, in view of (2.67), $\widehat{Y}_{\gamma, d}(\hat{v}) \leq \nu(\mu, \gamma, d) \leq (\mathbb{I} - \frac{1}{2}\widetilde{\alpha}_Y)(\widetilde{\delta})$. By the same line of reasoning as in the proof of Theorem 2.2, we deduce that there exist $\hat{\beta} \in \mathcal{KL}$, which is uniform in γ and d , such that any solution $\hat{\phi}$ to (2.50) initialized at $\sigma(x) \leq \Delta$ and any $k \in \mathbb{Z}_{\geq 0}$,

$$\widehat{Y}_{\gamma, d}(\hat{\phi}(k, x)) \leq \max\{\hat{\beta}(\widehat{Y}_{\gamma, d}(x), k), (\mathbb{I} - \frac{1}{2}\widetilde{\alpha}_Y)(\widetilde{\delta})\}. \quad (2.70)$$

^{†9} This holds with the same reasoning as (2.27) in the proof of Theorem 2.2, with the following rectification. In (2.27), we have utilized $\Upsilon(\cdot, \gamma, d) - (1 - \gamma)\mathbb{I} \in \mathcal{K}_\infty$ to show that $\gamma\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1} + \Upsilon(\cdot, \gamma, d) \in \mathcal{K}_\infty$. Here, it suffices to note that $\widehat{\Upsilon}(\cdot, \mu, \gamma, d) - (1 - \gamma)\mathbb{I}$ is non-decreasing, henceforth by following the same steps as in (2.27) the mentioned bound is found.

Finally, recall $\underline{\alpha}_Y(\sigma(x)) \leq \widehat{Y}_{\gamma,d}(x) \leq \overline{\alpha}_Y(\sigma(x)) + \eta(\epsilon, \gamma, d) \leq 2 \max\{\overline{\alpha}_Y(\sigma(x)), \eta(\epsilon, \gamma, d)\}$ and $\widetilde{\delta} = (\mathbb{I} - \frac{1}{2}\widetilde{\alpha}_Y)^{-1} \circ \underline{\alpha}_Y(\delta)$. Then, since $\widehat{\beta} \in \mathcal{KL}$, we derive

$$\begin{aligned} \sigma(\widehat{\phi}(k, x)) &\leq \max\{\underline{\alpha}_Y^{-1}\left(\widehat{\beta}(2\overline{\alpha}_Y(\sigma(x)), k)\right), \\ &\quad \underline{\alpha}_Y^{-1}\left(\widehat{\beta}(2\eta(\epsilon^*, \gamma, d), 0)\right), \delta\}. \end{aligned} \quad (2.71)$$

Let $\epsilon^* \in [0, \epsilon^*_{(2.69)})$ be such that $\underline{\alpha}_Y^{-1}\left(\widehat{\beta}(2\eta(\epsilon^*, \gamma, d), 0)\right) \leq \delta$, we obtain

$$\sigma(\widehat{\phi}(k, x)) \leq \max\{\underline{\alpha}_Y^{-1}\left(\widehat{\beta}(2\overline{\alpha}_Y(\sigma(x)), k)\right), \delta\}. \quad (2.72)$$

Thus (2.51) holds with $\beta(s, k) = \underline{\alpha}_Y^{-1}\left(\widehat{\beta}(2\overline{\alpha}_Y(s), k)\right)$.

We now prove the last part of Theorem 4. We have already noted the existence of a pair (γ^*, d^*) such that inequalities (2.66)-(2.68) hold for any $\gamma \in (\gamma^*, 1]$ and $d \in (d^*, \infty]$ for fixed μ . Assume, without loss of generality, that $2 \frac{\eta(\epsilon^*_{(2.69)}, \gamma^*, d^*)}{\gamma^*} \leq \mu$. It follows that $\eta(\epsilon, \gamma, d) \leq \eta(\epsilon^*_{(2.69)}, \gamma^*, d^*)$, as η is assumed here to be non-decreasing in γ and d , and is of class \mathcal{K}_∞ in ϵ . Since $\gamma^* \leq \gamma$, $\frac{1}{\gamma} \leq \frac{1}{\gamma^*}$. Hence, $2 \frac{\eta(\epsilon, \gamma, d)}{\gamma} \leq \mu$. Therefore, the remaining inequality (2.69) is verified for any $(\epsilon, \gamma, d) \in [0, \epsilon^*) \times \mathcal{P}_{\gamma^*, d^*}$ and (2.71) holds. All that is left to find is $\epsilon^* \leq \epsilon^*_{(2.69)}$ such that $\underline{\alpha}_Y^{-1}\left(\widehat{\beta}(2(\epsilon^*, \gamma^*, d^*), 0)\right) \leq \delta$, hence $\underline{\alpha}_Y^{-1}\left(\widehat{\beta}(2\eta(\epsilon, \gamma, d), 0)\right) \leq \delta$ and the proof is complete. \blacksquare

Theorem 2.4 ensures a semiglobal practical stability property where the adjustable parameters are not only γ, d as before, but also ϵ . We stress that ϵ^* , the upper-bound on ϵ , has to be chosen as a function of (γ, d) in general, not (γ^*, d^*) . This is due to the fact that the error term $\eta(\epsilon, \gamma, d)$ is potentially increasing and unbounded in γ and d . This is the case for instance when $\eta(\epsilon, \gamma, d) = \frac{1-\gamma^d}{1-\gamma}\epsilon$, which explodes as $(\gamma, d) \rightarrow (1, \infty)$ as already mentioned. We are thus forced to adapt ϵ to (γ, d) . This is no longer the case when $\eta(\epsilon, \gamma, d)$ is non-increasing in γ and d , as stated in the last part of Theorem 2.4.

While the determination of the pair (γ^*, d^*) such that (2.51) holds is as difficult as in Theorem 2.2, the next lemma states that any pair (γ^*, d^*) such that (2.51) holds for given (δ, Δ) , ensures that (2.23) holds with tuple $(\gamma^*, d^*, \delta, \Delta)$, but the reverse may not be true.

Lemma 2.5. *Let $(\gamma^*, d^*, \delta, \Delta)$ be such that Theorem 2.4 holds for some η and ϵ^* . Then Theorem 2.2 is verified with the same tuple $(\gamma^*, d^*, \delta, \Delta)$. \square*

Proof. Let $s \in \mathbb{R}_{\geq 0}$, $\gamma \in (0, 1]$, $d \in \mathbb{Z}_{>0} \cup \{\infty\}$, $\delta, \Delta > 0$ and recall the definitions for Υ , $\widehat{\Upsilon}$ and respective $\widetilde{\alpha}_Y$ from the proofs of Theorem 2.1 and Theorem 2.4. It follows that $\Upsilon(s, \gamma, d) = (1 - \gamma)s + \gamma^d \bar{\alpha}_V \circ \alpha_W^{-1} \left(\frac{\mathbb{I} - \alpha_W \circ (\bar{\alpha}_V + \bar{\alpha}_W)^{-1}}{\gamma} \right)^{(d)}(s) \leq (1 - \gamma)s + \gamma^d \bar{\alpha}_V \circ \alpha_W^{-1} \left(\frac{\mathbb{I} - \alpha_W \circ (\bar{\alpha}_V + \bar{\alpha}_W)^{-1} \circ \frac{\mathbb{I}}{2}}{\gamma} \right)^{(d)}(s)$. On the other hand, for any $\mu \geq 0$, we have that $\left(\frac{\mathbb{I} - \alpha_W \circ (\bar{\alpha}_V + \bar{\alpha}_W)^{-1} \circ \frac{\mathbb{I}}{2}}{\gamma} \right)^{(d)} \leq \max \left\{ \left(\frac{\mathbb{I} - \alpha_W \circ (\bar{\alpha}_V + \bar{\alpha}_W)^{-1} \circ \frac{\mathbb{I}}{2}}{\gamma} \right)^{(d)}, \mu \right\}$. Therefore $\Upsilon(s, \gamma, d) \leq \widehat{\Upsilon}(s, \mu, \gamma, d) - \frac{\mu^2}{2} \leq \widehat{\Upsilon}(s, \mu, \gamma, d)$. On the other hand, we have $\widetilde{\alpha}_Y = \alpha_W \circ \bar{\alpha}_Y^{-1}$ in Theorem 2.1 and $\widetilde{\alpha}_Y = \alpha_W \circ \bar{\alpha}_Y^{-1} \circ \frac{\mathbb{I}}{2}$. Clearly, when (2.66) holds for pair (δ, Δ) for all $(\mu, \gamma, d) \in [0, \mu^*] \times (\gamma^*, 1] \times (d^*, \infty]$, it follows that $\Upsilon(s, \gamma, d) \leq \widehat{\Upsilon}(s, \mu, \gamma, d) \leq (1 - \frac{\gamma}{2}) \alpha_W \circ \bar{\alpha}_Y^{-1} \circ \frac{\mathbb{I}}{2} \leq (1 - \frac{\gamma}{2}) \alpha_W \circ \bar{\alpha}_Y^{-1}(s)$, thus equation (2.24) holds for all $(\gamma, d) \in (\gamma^*, 1] \times (d^*, \infty]$ and the proof is complete. \blacksquare

Lemma 2.5 is useful in practice to have lower bounds on γ^* and d^* when the considered algorithm produces near-optimal inputs according to Assumption 2.3, as illustrated in the next section.

Regarding optimality, the following corollary provides a relationship between the infinite-horizon discounted optimal cost $V_{\gamma, \infty}(x)$ and $\widehat{V}_{\gamma, d}(x)$.

Corollary 2.4. *Suppose that Assumption 2.1 and 2.3 hold, let $\epsilon \geq 0$, $(\gamma, d) \in \mathcal{P}$ and $x \in \mathbb{R}^n$ and consider $v_{\gamma, d}(x)$ defined in Theorem 2.3. The following holds.*

$$\widehat{V}_{\gamma, d}(x) - \eta(\epsilon, \gamma, d) \leq V_{\gamma, \infty}(x) \leq \widehat{V}_{\gamma, d}(x) + \gamma^d v_{\gamma, d}(x). \quad (2.73)$$

□

Proof. The proof immediately follows from Assumption 2.3 and item (i) of Theorem 2.3, which holds from Assumption 2.1. In particular $\widehat{V}_{\gamma, d}(x) - \eta(\epsilon, \gamma, d) \leq V_{\gamma, d}(x) \leq V_{\gamma, \infty}(x) \leq V_{\gamma, d}(x) + \gamma^d v_{\gamma, d}(x) \leq \widehat{V}_{\gamma, d}(x) + \gamma^d v_{\gamma, d}(x)$.

2.6 Illustrative example: an inverted pendulum

We consider the model of an inverted pendulum discretized by Euler scheme with sampling period $T > 0$,

$$\begin{aligned} x_1^+ &= x_1 + Tx_2 \\ x_2^+ &= x_2 + T(a \sin(x_1) - bx_2 + cu) \end{aligned} \quad (2.74)$$

where $x_1 \in \mathbb{R}$ is the angular position of the pendulum, with $x_1 = 0$ being the upper-position, $x_2 \in \mathbb{R}$ the angular velocity and $u \in \mathbb{R}$ is a controllable torque at the rotation axis. The constants $a, b, c > 0$ are related to the mass, the dissipation and the motor gain, respectively. We have $(x_1, x_2) = x \in \mathbb{R}^2$ and $u \in \mathcal{U}(x) = \mathbb{R}$. Let $\sigma(x) = |x_1| + |x_2|$ and consider cost (2.2) with $\ell(x, u) = \sigma(x) + r|u|$ for some $r > 0$, for any $(x, u) \in \mathbb{R}^2 \times \mathbb{R}$.

First, we verify that SA holds by applying Theorem 1 and item d_3) of Theorem 2 in [39]. We have that items a) to c) of Theorem 1 in [39] hold. Item d_3) of Theorem 2 in [39] is satisfied since $|u| \leq \ell(x, u)$, thus $|u| \rightarrow \infty$, $\ell(x, u) \rightarrow \infty$ for any $x \in \mathbb{R}^2$. For item e) of Theorem 1 in [39], let $x \in \mathbb{R}^2$ and consider the infinite sequence

$$v(x) := \left\{ \frac{1}{c}(-a \sin(x_1) + bx_2 - \frac{x_1 + Tx_2}{T^2}), \frac{1}{c}(-a \sin(x_1 + Tx_2) + b\frac{x_1 + Tx_2}{T}), 0, \dots \right\}.$$

It follows that $\phi(1, x, v(x)|_1) = (x_1 + Tx_2, -\frac{x_1 + Tx_2}{T})$, and $\phi(k, x, v(x)|_k) = 0$ for $k \geq 2$. For any $\gamma \in (0, 1]$, we have

$$\begin{aligned} J_{\gamma, \infty}(x, v(x)) &= |x_1| + |x_2| + r \left| \frac{1}{c}(-a \sin(x_1) + bx_2 - \frac{x_1 + Tx_2}{T^2}) \right| \\ &\quad + \gamma(|x_1 + Tx_2| + |\frac{x_1}{T} + x_2|) \\ &\quad + r \left| \frac{1}{c}(-a \sin(x_1 + Tx_2) + b\frac{x_1 + Tx_2}{T}) \right| \\ &\leq |x_1| + |x_2| + \frac{r}{c}(a|x_1| + b|x_2| + \frac{1}{T^2}|x_1| + \frac{1}{T}|x_2|) \\ &\quad + \gamma(|x_1| + T|x_2| + \frac{1}{T}|x_1| + |x_2|) \\ &\quad + \frac{r}{c}(a(|x_1| + T|x_2|) + b(\frac{1}{T}|x_1| + |x_2|)) \\ &\leq \theta_1(\gamma, r, T)|x_1| + \theta_2(\gamma, r, T)|x_2|, \end{aligned}$$

with $\theta_1(\gamma, r, T) := 1 + \frac{ra}{c} + \frac{r}{cT^2} + \gamma(1 + \frac{1}{T} + \frac{ra}{c} + \frac{rb}{cT})$ and $\theta_2(\gamma, r, T) := 1 + \frac{rb}{c} + \frac{r}{cT} + \gamma(T + 1 + \frac{rT}{c} + \frac{rb}{c})$, which is finite for all $T > 0$. Thus, Theorem 2 of [39] is verified and SA holds. Furthermore, since $J_{1,\infty}(x, v(x)) \leq \theta_1(1, r, T)|x_1| + \theta_2(1, r, T)|x_2|$ we verify item (i) of Assumption 2.1 is satisfied with $\bar{\alpha}_V = \max\{\theta_1(1, r, T), \theta_2(1, r, T)\}\mathbb{I}$. From $\ell(x, u) \geq \ell(x, 0) = \sigma(x)$, item (ii) of Assumption 2.1 is verified with $W = \bar{\alpha}_W = 0$ and $\alpha_W = \mathbb{I}$.

In the following, we take $a = b = c = 1$, with time-step $T = 1$ and $r = 1$. In this case, $\bar{\alpha}_V = 7\mathbb{I}$. In view of Corollary 2.2, we find $\bar{d} = \left\lfloor \frac{0 - \ln 7^2}{\ln 6 - \ln 7} \right\rfloor = 25$ and $\bar{\gamma} = \frac{6}{7}$, thus global exponential stability is guaranteed for any pair (γ^*, d^*) with $\gamma^* \geq \bar{\gamma}$ and $d^* \geq \bar{d}$ when the optimal sequence of inputs is applied to system (2.74). This analysis is of course conservative, and a different derivation of $\bar{\alpha}_V$ might provide different bounds on (γ^*, d^*) .

Because we do not know how to compute optimal sequences of inputs, we use a scheme based on AVI to generate the inputs. This scheme relies on a simple finite difference approximation, with $N = 33^2$ points equally distributed in $[-\pi, \pi] \times [-\pi, \pi]$ for the state space, and 101 quantized inputs in $[-10, 10]$ centered at 0. As a result, Assumption 2.3 holds for some η , which can be derived by adapting the results from [16] for aggregation. Per Theorem 2.4, system (2.74) controlled by such scheme satisfies (2.51). Figure 2.1 provides plots of $\sigma(x)$ for the initial condition $(3, 0)$ for different values of (γ, d) . As expected, when γ or d are too small, $\sigma(\hat{\phi}(\cdot, x))$ does not converge to a “small” neighborhood of the origin, where we recall that $\hat{\phi}(\cdot, x)$ is the solution of (2.50) controlled by the near-optimal sequence.

To analyse the impact of (γ, d) on the closed-loop system, we consider three initial conditions, namely $(\pi, 0)$, $(\frac{\pi}{2}, \frac{\pi}{2})$ and $(\frac{\pi}{10}, 0.1)$, for different pairs (γ, d) and we have checked numerically whether the state measure converges to the set $\mathcal{S} := \{z \in \mathbb{R}^2 : \sigma(z) \leq 0.5\}$ in 11 steps, and that it stays in \mathcal{S} for at least 40 steps. The obtained results are summarized in Figure 2.2. As expected by Theorem 2.4, we see in Figure 2.2 that both (γ, d) have to be sufficiently large for $\sigma(\hat{\phi}(\cdot, x))$ to converge close to the origin. Furthermore, Theorem 2.4 suggests that stability may be lost if ϵ is not small enough for given (γ, d) , this is not apparent for our test case with $N = 33^2$. However, by reducing the number of points for the state space from $N = 33^2$ to $N = 31^2$, we observe a degradation of our convergence test

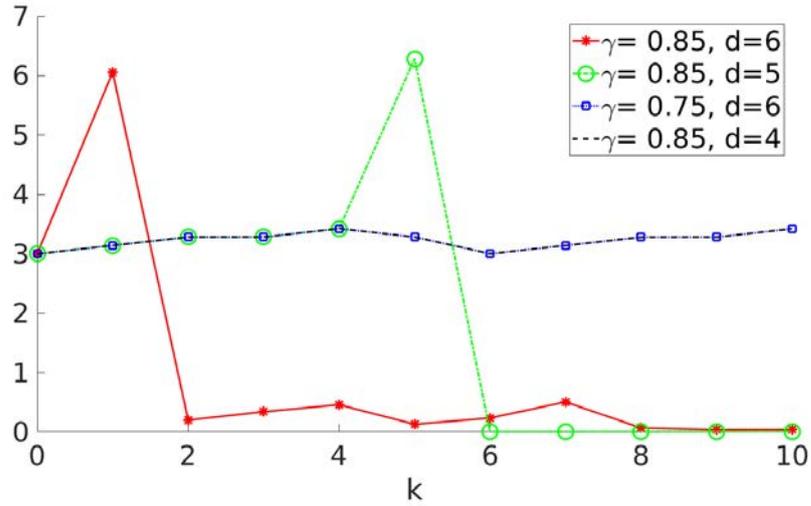


Figure 2.1: $\sigma(\hat{\phi}(\cdot, x))$ for 4 different pairs (γ, d) and $x = (3, 0)$.

for γ to close to 1 and d too large, see Figure 2.3. This further illustrates the statement made in Theorem 2.4 in extension to those made in Theorem 2.2. Namely, that pair (γ, d) cannot be simply taken close to $(1, \infty)$, as they have a role in addition to ϵ in keeping the error term η in check.

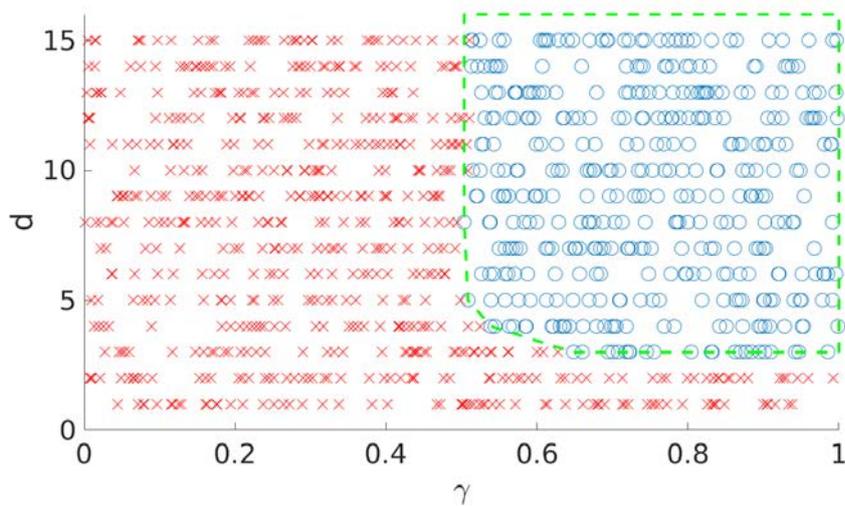


Figure 2.2: Convergence testing of 1000 sample pairs (γ, d) , $N = 33^2$. Symbol \circ denotes convergence to \mathcal{S} , while \times the converse.

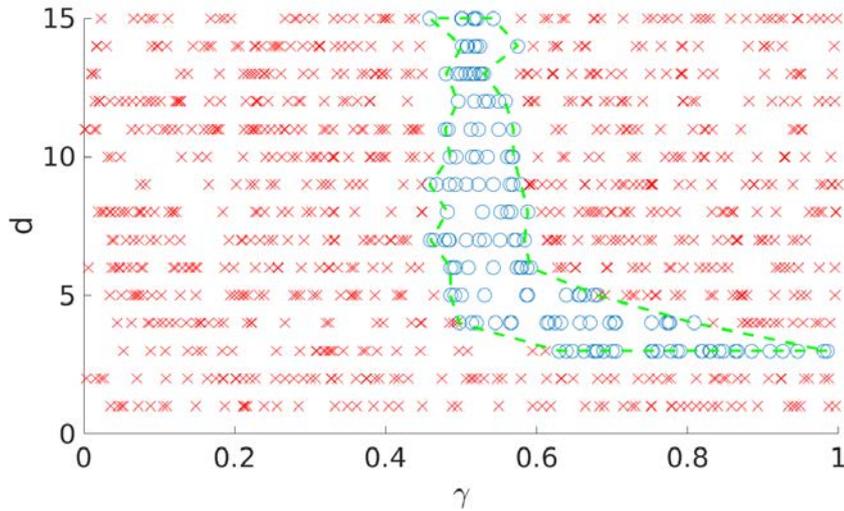


Figure 2.3: Convergence testing of 1000 sample pairs (γ, d) , $N = 31^2$. Symbol \circ denotes convergence to \mathcal{S} , while \times the converse.

2.7 Conclusion

We have analyzed the stability of nonlinear discrete-time systems controlled by a sequence of inputs that minimizes a discounted finite-horizon cost and is implemented in a receding-horizon fashion. In general, uniform semiglobal practical stability is ensured under suitable stabilizability and detectability conditions. Under additional assumptions, semiglobal asymptotic stability and uniform global exponential stability properties are guaranteed. The Lyapunov function used to prove stability is shown to be continuous under extra assumptions, hence endowing stability with some nominal robustness. We compared our results with previous works of the literature on undiscounted finite-horizon and discounted infinite-horizon optimal control, respectively. We have identified conditions under which the bounds we provide on discount factor and on horizon length are less conservative than those in [55] and [24]. Furthermore, we provide new relationships between the optimal value functions of the discounted, undiscounted, infinite-horizon, finite-horizon costs, which differ from those typically found in the (A)DP literature. Finally, we have analyzed stability when the available inputs are only near-optimal, for which an illustrative example was provided.

Given our motivations underlined in Chapter 1, the results of this chapter provide the stability guarantees for a large class of nonlinear optimal control algorithms. We have even provided novel bounds on the optimality applicable for these methods. One point remains to be seen: whether stability can be exploited at the core of an algorithm, for its benefit. This is the underlying theme of the next chapter.

Chapter 3. Control of switched nonlinear discrete-time systems: a planning-based approach

3.1	Introduction	53
3.2	Problem statement	54
3.3	OP_{min}	55
3.4	Main results	65
3.5	Example: a cubic integrator	82
3.6	Case study	86
3.7	Conclusion	92

3.1 Introduction

Originating in the artificial intelligence literature, optimistic planning (OP) is an algorithm that generates near-optimal control inputs for generic nonlinear discrete-time systems whose input set is finite [32]. This technique is therefore relevant for the near-optimal control of nonlinear switched systems, for which the switching signal is the control. However, OP exhibits several limitations, which prevent its application in a standard control context. First, it requires the stage cost to take values in $[0,1]$, an unnatural prerequisite as it

excludes, for instance, quadratic stage costs. Second, it requires the cost function to be discounted. Third, it applies for reward maximization, and not cost minimization.

In this chapter, we modify OP to overcome these limitations, and we call the new algorithm OP_{\min} , which comes in two flavors: $\text{OP}_{\min, \text{budget}}$ and $\text{OP}_{\min, \text{stop}}$. $\text{OP}_{\min, \text{budget}}$ has a fixed computational budget of size B , which has to be sufficiently large for our purpose as explained in the following. On the other hand, $\text{OP}_{\min, \text{stop}}$ is centered around a stopping criterion, that is designed by the user so that the induced computational budget adapts to the desired stability property and the current state.

We then make stabilizability and detectability assumptions as in Chapter 2, and derive near-optimality guarantees for OP_{\min} and we show that the obtained bound has major advantages compared to the bound originally given by OP. In addition, we prove that a system whose inputs are generated by OP_{\min} in a receding-horizon fashion exhibits stability properties, given an appropriate choice of budget B or stopping criterion. The results follow similar proof techniques as those in Chapter 2, except that here the horizon of the cost function varies with the state as we will see, which requires some adjustments. In addition, as a result, OP_{\min} provides a new tool for the near-optimal, stable control of nonlinear switched discrete-time systems for generic cost functions. We illustrate the potential of OP_{\min} on a problem where we want to optimally switch between given feedback laws.

The results for $\text{OP}_{\min, \text{budget}}$ corresponds to [23]. The results for $\text{OP}_{\min, \text{stop}}$ will be the subject to a journal publication.

3.2 Problem statement

Consider the system

$$x^+ = f_u(x), \tag{3.1}$$

with state $x \in \mathbb{R}^n$, control input $u \in \mathcal{U}$ where $\mathcal{U} := \{1, \dots, M\}$ is a finite set of admissible inputs with $M \geq 2$, and $f_u : \mathbb{R}^n \rightarrow \mathbb{R}^n$ for every $u \in \mathcal{U}$. We use $\phi(k, x, \mathbf{u}|_k)$ to denote the solution to system (3.1) at time $k \in \mathbb{Z}_{\geq 0}$ with initial condition x and inputs sequence $\mathbf{u}|_k = [u_0, u_1, \dots, u_{k-1}]$, with the convention $\phi(0, x, \cdot) = \phi(0, x, \emptyset) = x$.

We consider the infinite-horizon cost

$$J_\infty(x, \mathbf{u}) := \sum_{k=0}^{\infty} \ell_{u_k}(\phi(k, x, \mathbf{u}|_k)), \quad (3.2)$$

where $x \in \mathbb{R}^n$ is the initial state, \mathbf{u} is an infinite sequence of admissible inputs, $\ell_u : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$ is the stage cost given input $u \in \mathcal{U}$. Finding an infinite sequence of inputs which minimizes (3.2) given $x \in \mathbb{R}^n$ is very difficult in general, as the particular case of linear switched systems with quadratic stage cost already shows [44, 79]. We therefore aim at generating sequences of inputs that *nearly* minimize (3.2) instead, in a sense made precise in the following. For this purpose, we revise optimistic planning (OP) as originally developed in [32]. We call these new algorithms $OP_{min,budget}$ and $OP_{min,stop}$. Furthermore, we aim at ensuring stability properties for the induced closed-loop system. The algorithm is presented in the next section.

Remark 3.1. $OP_{min,budget}$ results in [23] accepts a discount factor $\gamma \in (0, 1]$, like in Chapter 2. We opt to drop the discount factor in this chapter for the sake of consistency between the two versions of OP_{min} . \square

3.3 OP_{min}

3.3.1 Main idea

Both algorithms that we present evaluates *finite*-horizon costs given any initial state $x \in \mathbb{R}^n$

$$J_d(x, \mathbf{u}_d) := \sum_{k=0}^d \ell_{u_k}(\phi(k, x, \mathbf{u}_d|_k)), \quad (3.3)$$

where $d \in \mathbb{Z}_{>0}$ is a horizon, and $\mathbf{u}_d = [u_0, u_1, \dots, u_d] \in \mathcal{U}^{d+1}$. OP_{min} searches for optimal input sequences which minimizes exactly cost (3.3), given state x , that is

$$V_d(x) := \min_{\mathbf{u}_d} J_d(x, \mathbf{u}_d), \quad (3.4)$$

for some finite horizon $d \in \mathbb{Z}_{>0}$ and we denote \mathbf{u}_d^* a corresponding optimal input sequence of length d , which may be non-unique. Hence, $V_d(x) = J_d(x, \mathbf{u}_d^*(x))$. The horizon d in (3.4) is selected by the algorithm itself. In particular, both algorithms iteratively increase

the horizon d in (3.3) up to a horizon $d(x) \in \mathbb{Z}_{>0}$, which depends on the initial state x . For $\text{OP}_{\min, \text{budget}}$, $d(x)$ is the largest d finite-horizon optimal cost found with a fixed computational budget B , as defined precisely in Section 3.3.3. For $\text{OP}_{\min, \text{stop}}$, $d(x)$ is the first d finite-horizon optimal cost that verifies a stopping criterion as explained in more details in the following.

Remark 3.2. One might wonder why we do not present a $\text{OP}_{\min, \text{horizon } d}$, i.e. OP_{\min} that stops at a fixed horizon d , like in Chapter 2. To begin with, stability and optimality guarantees follows immediately for a fixed horizon from Chapter 2, hence it is of no particular technical detail. Nonetheless, we also believe that both $\text{OP}_{\min, \text{budget}}$ and $\text{OP}_{\min, \text{stop}}$ supersede such algorithm. $\text{OP}_{\min, \text{budget}}$ has the advantage of having a fixed computational budget, which might find sequences for larger horizons without sacrificing execution time. A fixed horizon has no such run-time guarantees. On the other hand, one might find a stopping criterion such that $\text{OP}_{\min, \text{stop}}$ has $d(x) = d$ for all $x \in \mathbb{R}^n$, hence being equivalent to a $\text{OP}_{\min, \text{horizon } d}$. However, as we will see, the guarantees for $\text{OP}_{\min, \text{stop}}$ are less conservative by forgoing a fixed horizon, and doing so would be needlessly computational intensive. Therefore, in the scope of this chapter, we see no clear benefit for such $\text{OP}_{\min, \text{horizon } d}$. \square

3.3.1.1 $\text{OP}_{\min, \text{budget}}$

For $\text{OP}_{\min, \text{budget}}$ the horizon $d(x)$ is simply the longest finite-horizon sequence of inputs found that exhausts a given budget B of computational results. As we will show, by increasing the budget B , we control the horizon $d(x)$, which will decrease the mismatch between $V_{d(x)}(x)$ and the optimal value function associated to cost (3.2) at x .

Given any $x \in \mathbb{R}^n$ and budget $B > 0$, the cost function that $\text{OP}_{\min, \text{budget}}$ explicitly calculates is denoted by

$$V_B(x) := J_{d(x)}(x, \mathbf{u}_{d(x)}(x)). \quad (3.5)$$

We use the notation V_B in (3.5), instead of $V_{d(x)}$, to emphasize that the returned cost function is parameterized by B . Problem (3.5) is tractable when $d(x)$ is finite, as the input set \mathcal{U} is finite. In this case, a brute-force approach can solve it by developing all possible sequences. However, this is computationally intensive, in particular when $d(x)$

is large, as the computational cost grows exponentially with the horizon. $OP_{min,budget}$ instead intelligently explores the possible sequences to solve (3.5) with potentially larger horizons with the same computation budget compared to a brute-force approach [32]. In the following section, we show that longer horizons imply smaller near-optimality bounds, and are therefore desirable.

3.3.1.2 $OP_{min,stop}$

For $OP_{min,stop}$ we select horizon $d(x)$ implicitly such that

$$\sigma(\phi(d(x), x, \mathbf{u}_{d(x)}^*(x)|_{d(x)})) \leq c_{stop}(\varepsilon, x),$$

where

- $\phi(d(x), x, \mathbf{u}_{d(x)}^*(x)|_{d(x)})$ is the last state reached by applying the finite-horizon sequence $\mathbf{u}_{d(x)}^*(x)$.
- $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$ is a “measuring” function that we use to define stability as in Chapter 2, which depends on the problem.
- $c_{stop}(\varepsilon, x) \geq 0$ is a stopping criterion which we design and which may depend on state vector x and a vector of tunable parameters $\varepsilon \in \mathbb{R}^{n_\varepsilon}$, with $n_\varepsilon > 0$. The design of c_{stop} is explained in the following.

Hence

$$d(x) := \min \left\{ d \in \mathbb{Z}_{>0} : \sigma(\phi(d, x, \mathbf{u}_d^*(x)|_d)) \leq c_{stop}(\varepsilon, x) \right\}. \quad (3.6)$$

As we will show, by controlling the “size” of the last state $\phi(d(x), x, \mathbf{u}_{d(x)}^*(x)|_{d(x)})$ through function c_{stop} (and parameter ε), we control directly, for each $x \in \mathbb{R}^n$, the mismatch between $V_{d(x)}(x)$ and the optimal value function associated to cost (3.2) at x .

Altogether, given any $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}^{n_\varepsilon}$, the cost function that OP_{min} explicitly calculates is denoted by

$$V_\varepsilon(x) := J_{d(x)}(x, \mathbf{u}_{d(x)}(x)). \quad (3.7)$$

We use the notation V_ε in (3.7), instead of $V_{d(x)}$, to emphasize that the returned cost function is parameterized by ε . Problem (3.7) is tractable when $d(x)$ is finite, for the same

reasons as $\text{OP}_{\min, \text{budget}}$. However, for $d(x)$ in (3.6) to be finite for any $x \in \mathbb{R}^n$, we need to make assumptions.

Remark 3.3. Given the similarities of both algorithms, we will only provide proofs for $\text{OP}_{\min, \text{stop}}$ in this thesis. Some proofs for $\text{OP}_{\min, \text{budget}}$ can be found in [23]. \square

3.3.2 Assumptions

We first assume that the optimization goal (3.2) is well-posed in the following sense as in Section 2.2.

Standing Assumption 1 (SA1). For any $x \in \mathbb{R}^n$, there exists an infinite sequence of admissible inputs $\mathbf{u}_\infty^*(x)$, called *optimal input sequence*, which minimizes (3.2), i.e. $V_\infty(x) := J_\infty(x, \mathbf{u}_\infty^*(x))$ is finite. \square

General conditions to ensure SA1 can be found in [39] as mentioned in Chapter 3.4.

We also make the next general stabilizability and detectability assumptions on system (3.1), stage cost ℓ , as in Chapter 2.

Standing Assumption 2 (SA2). There exist $\bar{\alpha}_V, \alpha_W \in \mathcal{K}_\infty$, continuous functions $W, \sigma : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$, $\bar{\alpha}_W : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ continuous, non-decreasing and zero at zero, such that the following conditions hold.

(i) For any $x \in \mathbb{R}^n$,

$$V_\infty(x) \leq \bar{\alpha}_V(\sigma(x)). \quad (3.8)$$

(ii) For any $x \in \mathbb{R}^n$, $u \in \mathcal{U}$,

$$W(x) \leq \bar{\alpha}_W(\sigma(x)) \quad (3.9)$$

$$W(f_u(x)) - W(x) \leq -\alpha_W(\sigma(x)) + \ell_u(x). \quad (3.10)$$

\square

To recall, item (i) is related to the stabilizability of system (3.1) with respect to stage cost ℓ_u in relation to σ , and item (ii) of SA2 is a detectability property of the stage cost ℓ_u with respect to σ .

We need an extra assumption only for $OP_{min,stop}$ on c_{stop} , which is made without loss of generality as we design c_{stop} .

Standing Assumption 3 (SA3). For any $\varepsilon \in \mathbb{R}^{n_\varepsilon}$ and any $x \in \{z \in \mathbb{R}^n : \sigma(z) > 0\}$, $c_{stop}(\varepsilon, x) > 0$. \square

SA3 formalizes which stopping functions c_{stop} guarantee that $OP_{min,stop}$ terminates. Possible candidate functions are, e.g., $c_{stop}(\varepsilon, x) = \alpha(\sigma(x))$ for some $\alpha \in \mathcal{K}$ in which case there is no parameter ε , $c_{stop}(\varepsilon, x) = \varepsilon\alpha(\sigma(x))$, $c_{stop}(\varepsilon, x) = \varepsilon$, with $\varepsilon > 0$, or combinations like $c_{stop}(\varepsilon, x) = \max\{\varepsilon_1\sigma(x), \varepsilon_2\}$ for $\varepsilon = (\varepsilon_1, \varepsilon_2)$ and $x \in \mathbb{R}^n$. We stress that c_{stop} is not required to be positive definite, i.e. we accept $c_{stop}(\varepsilon, x) > 0$ for $\sigma(x) = 0$. By shaping the terminating function c_{stop} , we can tighten (or relax) near-optimality properties as shown in the following. The next statement ensures that, given any $x \in \mathbb{R}^n$, $d(x)$ as defined in (3.6) is finite.

Proposition 3.1. For all $x \in \mathbb{R}^n$, $d(x)$ in (3.6) is finite and $V_\varepsilon(x)$ in (3.7) is thus well-defined. \square

Proof. Let $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}^{n_\varepsilon}$. In view of Lemma B.1 in the appendix, we distinguish two cases. When $\sigma(x) = 0$, it follows that $\sigma(\phi(d, x, \mathbf{u}_d^*(x))) \leq \alpha_W^{-1} \circ (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^d \circ \bar{\alpha}_Y(0) = 0$ for any $d \in \mathbb{Z}_{>0}$ according to item (i) of Lemma B.1. Hence, $\sigma(\phi(d, x, \mathbf{u}_d^*(x))) \leq c_{stop}(\varepsilon, x)$ for any $d \in \mathbb{Z}_{>0}$ since $c_{stop}(\varepsilon, x) \geq 0$. When $\sigma(x) > 0$, $\sigma(\phi(d, x, \mathbf{u}_d^*(x))) \leq \alpha_W^{-1} \circ (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^d \circ \bar{\alpha}_Y(\sigma(x))$ and this upper-bound can be made arbitrarily close to 0 by increasing d , according to item (iii) of Lemma B.1. Hence, there exists a finite d sufficiently large such that $\sigma(\phi(d, x, \mathbf{u}_d^*(x))) \leq \alpha_W^{-1}((\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^d(\bar{\alpha}_Y(\sigma(x)))) \leq c_{stop}(\varepsilon, x)$ as $c_{stop}(\varepsilon, x) > 0$ for $\sigma(x) > 0$. Thus, for any $x \in \mathbb{R}^n$ there is a finite $d \in \mathbb{Z}_{>0}$ such that $\sigma(\phi(d, x, \mathbf{u}_d^*(x))) \leq c_{stop}(\varepsilon, x)$. As a consequence, $d(x)$ in (3.6) is finite and so is V_ε in (3.7). \blacksquare

Remark 3.4. Note that $OP_{min,budget}$ requires no analogue result, as it will stop whenever the budget is exhausted, hence termination is guaranteed. \square

3.3.3 Algorithm description

Both OP_{\min} algorithms explore the possible choices of inputs *optimistically* until a stopping condition is verified. The computational resources utilized for this purpose are denoted as a budget $B \in \mathbb{Z}_{>0}$, which corresponds to $B + 1$ “leaf expansions”, and which is either fixed or adapted to the state x depending on the version of OP_{\min} . We denote by \mathcal{T} the exploration tree from initial state $x \in \mathbb{R}^n$, constructed from admissible input sequences and their respective cost (3.3). A leaf is a node of \mathcal{T} with no children, and the set of all leaves of \mathcal{T} is denoted $\mathcal{L}(\mathcal{T})$. At iteration $i \in \mathbb{Z}_{\geq 0}$, a leaf $L_i \in \mathcal{L}(\mathcal{T})$ is fully expanded. That is, for every $u \in \mathcal{U}$, we add a child to L_i labeled by the resulting state $f_u(L_i)$. We denote with a slight abuse of notation $\mathbf{u}(L_i)$ the input sequence from the root x to the state of leaf L_i . We also denote by $J(L_i) := J_{d(i)}(x, \mathbf{u}(L_i))$ cost (3.3) given by the sequence that takes x to the state of leaf L_i , with $d(i) := \text{depth}(L_i) - 1$, where $\text{depth}(\cdot)$ is the number of edges (or inputs) from the root to L_i . The leaf L_i the algorithm expands is the leaf with minimal associated cost $J(L)$ of all non-expanded leaves $L \in \mathcal{T}$. $\text{OP}_{\min, \text{budget}}$ simply terminates when the available budget is exhausted. $\text{OP}_{\min, \text{stop}}$, on the other hand, terminates when an optimal sequence candidate is found with $\sigma(L_i) \leq c_{\text{stop}}(\varepsilon, x)$, that is, verifies $\sigma(\phi(d, x, \mathbf{u}_d^*(x))) \leq c_{\text{stop}}(\varepsilon, x)$, see (3.6). This sequence exists according to Proposition 3.1. Both algorithms are formalized next.

Algorithm 1 Algorithm for $OP_{min,budget}$

Input: budget B , state x **Output:** depth explored $d(x)$, sequence $\mathbf{u}_{d(x)}^*(x)$, cost $V_B(x)$ *Initialization :*1: $d \leftarrow -1$ 2: tree $\mathcal{T} \leftarrow \{[], 0\}$ {the empty sequence and cost 0}*Optimistic exploration*3: **for** $i = 0$ to B **do**4: find optimistic leaf $L_i \in \arg \min_{L \in \mathcal{L}(\mathcal{T})} J(L)$ *Add to \mathcal{T} the children of L_i :*5: for each child c of L_i , $\mathcal{T} \leftarrow \mathcal{T} \cup \{\mathbf{u}(c), J(c)\}$ 6: **if** $d < \text{depth}(L_i) - 1$ **then** {first expansion at new depth is d -horizon optimal}*Leaf selection*7: $S \leftarrow L_i$ 8: $d \leftarrow \text{depth}(L_i) - 1$ 9: **end if**10: **end for**11: **return** $d(x) \leftarrow d$ and $\{\mathbf{u}_{d(x)}^*(x), V_B(x)\} \leftarrow S$

Algorithm 2 Algorithm for $\text{OP}_{\min, \text{stop}}$

Input: $c_{\text{stop}}(\varepsilon, \cdot)$, state x

Output: depth explored $d(x)$, sequence $\mathbf{u}_{d(x)}^*(x)$, cost $V_\varepsilon(x)$

Initialization :

1: $d, i \leftarrow -1$

2: tree $\mathcal{T} \leftarrow \{[], 0\}$ {the empty sequence and cost 0}

Optimistic exploration

3: **while** true **do**

4: $i = i + 1$

5: find optimistic leaf $L_i \in \arg \min_{L \in \mathcal{L}(\mathcal{T})} J(L)$

Add to \mathcal{T} the children of L_i :

6: for each child c of L_i , $\mathcal{T} \leftarrow \mathcal{T} \cup \{\mathbf{u}(c), J(c)\}$

7: **if** $d < \text{depth}(L_i) - 1$ **then** {first expansion at new depth is d -horizon optimal}

Leaf selection

8: $S \leftarrow L_i$

9: $d \leftarrow \text{depth}(L_i) - 1$

10: **if** $\sigma(L_i) \leq c_{\text{stop}}(\varepsilon, x)$ **then**

11: break

12: **end if**

13: **end if**

14: **end while**

15: $B \leftarrow i$

16: **return** $d(x) \leftarrow d$ and $\{\mathbf{u}_{d(x)}^*(x), V_\varepsilon(x)\} \leftarrow S$

Steps in green, lines 4-5 of Algorithm 1 and lines 5-6 of Algorithm 2, are the optimistic exploration. Note that the expansion of the tree is independent from the “leaf selection” step, and is fully determined by the optimistic selection of leaves. Any sequence of inputs from descendants (children, children of children and so on) of a node N will have costs J greater than N , as $\ell_u(x) \geq 0$ for any $x \in \mathbb{R}^n$ and $u \in \mathcal{U}$. The optimistic choice then guarantees that $J(L_{i+1}) \geq J(L_i)$ for any iteration $i \in \mathbb{Z}_{\geq 0}$. This implies that the first leaf to be expanded at a depth $d' + 1$ will be a suitable candidate for $V_{d'}(x)$.

For $OP_{min,budget}$, we simply return the last found candidate calculated with budget B , and the output is $V_B(x)$. The proofs of the next two propositions follow from the same arguments as Proposition 3.4. Proposition 3.3 in particular follows simply by calculating the size of the shallowest possible tree of depth \bar{d} .

Proposition 3.2. *Given a budget $B \geq 1$, Algorithm 1 terminates with output $S = \{\mathbf{u}_{d(x)}^*(x), V_B(x)\}$ with horizon $d(x) \geq 0$. \square*

The horizon $d(x)$ in (3.5) depends on the given budget B , and will play a fundamental role in the near-optimality analysis provided later in Section 3.4. The next proposition provides a (conservative) relationship between budget B and a given lower bound on $d(x)$.

Proposition 3.3. *Given $\bar{d} \geq 0$ and budget $B \geq \frac{M^{\bar{d}+1}-1}{M-1}$, Algorithm 1 returns $S = \{\mathbf{u}_{d(x)}^*(x), V_B(x)\}$ with horizon $B - 1 \geq d(x) \geq \bar{d}$. \square*

Proposition 3.3 provides a relationship between a minimum desired horizon \bar{d} in (3.3) and the required budget to achieve it. This relationship is derived from the worst-case exploration, which happens when $OP_{min,budget}$ is forced to uniformly explore the possible choices of switches for a given horizon. Due to the optimistic exploration, for given $B = \frac{M^{\bar{d}+1}-1}{M-1}$, $d(x)$ is often much larger in practice than \bar{d} . In the original OP study [32], this fraction is quantified by means of the branching factor κ , which might be of interest for future work.

For $OP_{min,stop}$, the candidate $V_{d'}(x)$ is in turn tested for the terminating constraint (3.6). The output cost is $V_\varepsilon(x)$, corresponding to the first finite-horizon optimal input sequence that verifies (3.6), calculated with a varying budget B , which depends on x .

The terminating condition for Algorithm 2 is guaranteed to be eventually verified when Proposition 3.1 holds, as formalized in the next proposition.

Proposition 3.4. *Let $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}^{n\varepsilon}$. Algorithm 2 terminates with output $S = \{\mathbf{u}_{d(x)}^*(x), V_\varepsilon(x)\}$. \square*

Proof. Let $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}^{n\varepsilon}$. First, we show that S exactly calculates cost $V_{d'}(x)$ for some $d' \in \mathbb{Z}_{>0}$. The optimal property of output S to Algorithm 2 is fully determined in the particular iteration in which it is updated. Hence, let \mathcal{T}_i be the tree to be expanded at iteration $i \in \mathbb{Z}_{\geq 0}$, in which S is updated. We show now that the selected leaf S with cost $J(S)$, where $J(S)$ is the cost associated to leaf S , attains the optimum of horizon $d' := \text{depth}(S) - 1$, that is $J(S) = V_{d'}(x)$. Since $V_{d'}(x) \leq J(S)$ by the optimality of $V_{d'}(x)$, it suffices to prove $V_{d'}(x) \geq J(S)$. For this purpose, we proceed by contradiction, and we assume that $V_{d'}(x) < J(S)$. It follows from the fact that the input set \mathcal{U} is finite that a sequence that attains the optimum $V_{d'}(x)$ exists, i.e. there is a node $N \neq S$, decedent of root x and possibly not in \mathcal{T}_i , with cost $J(N) = V_{d'}(x)$. Since $\ell_u(x) \geq 0$ for any $x \in \mathbb{R}^n$ and $u \in \{1, \dots, M\}$, any ancestor (parents, parents of parents and so on) of N will have cost a lower cost than $J(N)$. Hence, let L'_i be the ancestor of N such that $L'_i \in \mathcal{L}(\mathcal{T}_i)$, thus $J(L'_i) \leq J(N)$. Then, we have $J(L'_i) \leq J(N) = V_{d'}(x) < J(S)$, that is $J(L'_i) < J(S)$. However, S is the optimistically chosen leaf $S = L_i$, and $J(S) = J(L_i) \leq J(L)$ for any leaf $L \in \mathcal{L}(\mathcal{T}_i)$, hence for leaf L'_i , it follows that $J(L'_i) < J(S) \leq J(L'_i)$: we have attained a contradiction. Therefore, $J(S) \leq V_{d'}(x)$ and since $V_{d'}(x) \leq J(S)$, we conclude $V_{d'}(x) = J(S)$. Thus, at every update of S , a new optimal sequence is found with increased horizon $d' \leftarrow d' + 1$. By Proposition 3.1, $d(x)$ is well-defined and there exists a sequence $\mathbf{u}_{d(x)}^*(x)$ such that $\sigma(\phi(d(x), x, \mathbf{u}_{d(x)}^*(x))) \leq c_{\text{stop}}(\varepsilon, x)$. In other words, $\sigma(L_i) \leq c_{\text{stop}}(\varepsilon, x)$ is bound to be verified in a finite number of expansion, and $d(x)$ as defined in (3.6) holds with $d(x) = \text{depth}(L_i) - 1$ holds. Therefore, Algorithm 2 is guaranteed to terminate, with outputs of $d(x)$ and $S = \{V_\varepsilon(x), \mathbf{u}_{d(x)}^*\}$ fully determined. \blacksquare

3.4 Main results

In this section, we analyze the near-optimality properties for both flavors of OP_{\min} . We then provide conditions under which system (3.1), whose inputs are generated in a receding-horizon fashion by a flavor of OP_{\min} , exhibits stability properties. Robustness of this stability property is ensured under mild regularity properties.

3.4.1 Relationship between OP_{\min} and $V_{\infty}(x)$

3.4.1.1 $\text{OP}_{\min, \text{budget}}$

The next theorem provides near-optimal properties for $\text{OP}_{\min, \text{budget}}$.

Theorem 3.1. *For any $x \in \mathbb{R}^n$ and $d(x) \in \mathbb{Z}_{>0}$,*

$$V_B(x) \leq V_{\infty}(x) \leq V_B(x) + v_B(x), \quad (3.11)$$

where $v_B(x) := \bar{\alpha}_V \circ \underline{\alpha}_Y^{-1} \circ (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^{(d(x))} \circ \bar{\alpha}_Y(\sigma(x))$, with $d(x)$ as returned by Algorithm 1. Here, $\alpha_Y = \underline{\alpha}_Y := \alpha_W, \bar{\alpha}_Y := \bar{\alpha}_V + \bar{\alpha}_W$ and $\alpha_W, \bar{\alpha}_W, \bar{\alpha}_V$ come from SA2. \square

The lower-bound in (3.11) trivially holds from the optimality of $V_{d(x)}(x)$ as $d(x) < \infty$, see Proposition 3.3. The upper-bound, on the other hand, implies that the infinite-horizon cost is at most $v_{d(x)}(x)$ away from the finite-horizon $V_B(x)$. The error term $v_B(x)$ has three desirable properties. First, when $\sigma(x)$ is small, so is $v_B(x)$. Second, $v_B(x)$ accepts the undiscounted case. Third, $v_B(x) \rightarrow 0$ when $B \rightarrow \infty$, as it implies $d(x) \rightarrow \infty$, which implies that $(\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^{(d(x))}(s) \rightarrow 0$ for any $s \geq 0$ per Lemma 2.4 from Chapter 2.

3.4.1.2 $\text{OP}_{\min, \text{stop}}$

The next theorem provides near-optimal properties for $\text{OP}_{\min, \text{stop}}$.

Theorem 3.2. *For any $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}^{n_{\varepsilon}}$,*

$$V_{\varepsilon}(x) \leq V_{\infty}(x) \leq V_{\varepsilon}(x) + v_{\varepsilon}(x), \quad (3.12)$$

where $v_{\varepsilon}(x) := \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x))$ with $\bar{\alpha}_V$ from SA2. \square

Proof. Let $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}^{n_\varepsilon}$, $d(x) \in \mathbb{Z}_{>0}$ as in (3.6), optimal sequence $[u_0^*, u_1^*, \dots, u_{d(x)}^*] := \mathbf{u}_{d(x)}^*(x)$, cost $V_\varepsilon(x)$ defined in (3.7) are well-formed by Proposition 3.1. Since V_ε is a finite-horizon optimal cost, $V_\varepsilon(x) \leq V_\infty(x)$. On the other hand, consider the infinite-horizon sequence $\mathbf{u} = [u_0^*, u_1^*, \dots, u_{d(x)-1}^*, \mathbf{u}_\infty^*(\phi(d(x), x, \mathbf{u}_{d(x)}^*(x)|_{d(x)})))]$ which exists by SA1. It follows from the optimality of $V_\infty(x)$ that $V_\infty(x) \leq J_\infty(x, \mathbf{u})$, and from the definition of \mathbf{u} that $J_\infty(x, \mathbf{u}) = V_\varepsilon(x) + V_\infty(\phi(d(x), x, \mathbf{u}_{d(x)}^*(x)|_{d(x)}))$, which is finite. By invoking item (i) of SA2, we derive $V_\infty(x) \leq V_\varepsilon(x) + \bar{\alpha}_V(\sigma(\phi(d(x), x, \mathbf{u}_{d(x)}^*(x)|_{d(x)})))$, and, by the definition of $d(x)$ in (3.6), $V_\infty(x) \leq V_\varepsilon(x) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x))$. ■

The lower-bound in (3.12) trivially holds as in Theorem 3.5. The upper-bound, on the other hand, implies that the infinite-horizon cost is at most $v_\varepsilon(x)$ away from the finite-horizon $V_\varepsilon(x)$. The error term $v_\varepsilon(x)$ is small when so is $c_{\text{stop}}(\varepsilon, x)$ as $\bar{\alpha}_V \in \mathcal{K}_\infty$, which again we can tune. We can therefore make $V_\varepsilon(x)$ as close as desired to $V_\infty(x)$ by adjusting c_{stop} , the price to pay will be more computations.

Remark 3.5. Compared to the term given in [32], which we recall is $\frac{\gamma^{d(x)}}{1-\gamma}$ for a discount factor $\gamma \in (0, 1)$, v_B in (3.11) and v_ε in (3.12) are finite in the absence of discount factor, $\gamma = 1$. Moreover, as mentioned above, $c_{\text{stop}}(\varepsilon, x)$ small, so is $v_\varepsilon(x)$, which is not the case in [32]. Similarly for $v_B(x)$ for B large. By exploiting stabilizability and detectability properties in SA2, we have obtained an error bound that forfeits the assumption $\ell \in [0, 1]$, accepts the undiscounted case $\gamma = 1$, depends on either on the selected stopping condition c_{stop} or budget B , and is non-uniform in x . □

In Theorem 3.1, we see how $\text{OP}_{\text{min}, \text{budget}}$ and its fixed budget nature can approximate $V_\infty(x)$ via a finite-horizon optimal sequence, where the quality of the approximation adapts to the “hardness” of the problem. Furthermore, better approximations can be found by increasing budget B , which indirectly increases the state dependent horizon $d(x)$. Theorem 3.2, on the other hand relates $\text{OP}_{\text{min}, \text{stop}}$ approximation of $V_\infty(x)$ in terms of the stopping criterion. Hence, a bound on the quality of the approximation is known a priori, and the algorithm then simply utilizes the necessary computational budget to fulfill this criteria.

3.4.2 Stability

We now consider the scenario where (3.1) is controlled in a receding horizon fashion by of OP_{\min} . That is, at each time instant $k \in \mathbb{Z}_{\geq 0}$ with state $\phi_k \in \mathbb{R}^n$, the first element of the optimal sequence $\mathbf{u}_{d(\phi_k)}^*(\phi_k)$, is calculated by either Algorithm 1 or Algorithm 2, and then applied to system (3.1).

3.4.2.1 $\text{OP}_{\min, \text{budget}}$

$\text{OP}_{\min, \text{budget}}$ leads to closed-loop system

$$x^+ \in f_{\mathcal{U}_B^*(x)}(x) =: F_B^*(x), \quad (3.13)$$

where $f_{\mathcal{U}_B^*(x)}(x)$ is the set $\{f_u(x) : u \in \mathcal{U}_B^*(x)\}$, and $\mathcal{U}_B^*(x) := \{u_0 : \exists u_1, \dots, u_{d(x)} \in \mathcal{U} \text{ such that } V_{d(x)}(x) = J_{d(x)}(x, [u_0, \dots, u_{d(x)}])\}$ is the set of the first input of $d(x)$ -horizon optimal input sequences at x , with $d(x)$ given by Algorithm 1. We denote by $\phi(k, x)$, with some abuse of notation, a solution to (3.13) at time $k \in \mathbb{Z}_{\geq 0}$ with initial condition $x \in \mathbb{R}^n$.

The next theorem provides stability guarantees for system (3.13).

Theorem 3.3. *Consider system (3.13) and suppose Assumption 2.1 holds. There exists $\beta \in \mathcal{KL}$ such that for any $\delta, \Delta > 0$, there exist $\bar{d} \in \mathbb{Z}_{>0}$ such that for any budget $B \geq \frac{M^{\bar{d}+2}-1}{M-1}$, any $x \in \{z \in \mathbb{R}^n : \sigma(z) \leq \Delta\}$, any solution $\phi(\cdot, x)$ to system (3.13) satisfies, for all $k \in \mathbb{Z}_{\geq 0}$*

$$\sigma(\phi(k, x)) \leq \max\{\beta(\sigma(x), k), \delta\}. \quad (3.14)$$

□

Theorem 3.3 provides a semiglobal practical stability property for set $\{z : \sigma(z) = 0\}$, like Theorem 2.2 in Chapter 2.3.2. This implies that solutions to (3.13), with initial state x such that $\sigma(x) \leq \Delta$, where Δ is any given (arbitrarily large) strictly positive constant, will converge to the set $\{z : \sigma(z) \leq \delta\}$, where δ is any given (arbitrarily small) strictly positive constant, by taking a budget B sufficiently large. Note that we take budget $B \geq \frac{M^{\bar{d}+2}-1}{M-1}$ to guarantee $d(x) > \bar{d}$, by Proposition 3.3. While OP was used in various control problems [15, 17, 60], such stability properties were never been proved before.

As in Chapter 2.3.3, by strengthening SA2, we can prove a global exponential stability property.

Corollary 3.1. *Suppose there exist $\bar{a}_V, a_W > 0$, such that SA2 holds with $\bar{\alpha}_V \leq \bar{a}_V \cdot \mathbb{I}$, $\bar{\alpha}_W \leq \bar{a}_W \cdot \mathbb{I}$ and $\alpha_W \geq a_W \cdot \mathbb{I}$. Let \bar{d} be such that*

$$\frac{\bar{a}_V}{a_W} \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W} \right)^{\bar{d}} < \frac{a_W}{\bar{a}_V + \bar{a}_W}. \quad (3.15)$$

Then, there exist $\lambda > 0$, such that for any budget $B \geq \frac{M^{\bar{d}+2}-1}{M-1}$ and for any $x \in \mathbb{R}^n$, the solution $\phi(\cdot, x)$ to system (3.13), satisfies

$$\sigma(\phi(k, x)) \leq \frac{\bar{a}_V + \bar{a}_W}{a_W} \sigma(x) e^{-\lambda k} \quad (3.16)$$

for all $k \in \mathbb{Z}_{\geq 0}$. □

Corollary 3.1 ensures a uniform global exponential stability property of $\{x : \sigma(x) = 0\}$ for (3.13). Inequality (3.15) is always feasible for \bar{d} sufficiently large. Indeed, $\frac{\bar{a}_V}{a_W} \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W} \right)^{\bar{d}} \rightarrow 0$ when $\bar{d} \rightarrow \infty$, since^{†1} $\left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W} \right) \in (0, 1)$.

Remark 3.6. It is possible to relax Corollary 3.1 conditions and derive semiglobal asymptotic results, similarly to Corollary 2.2, or Corollary B.1 for $\text{OP}_{\min, \text{stop}}$. □

3.4.2.2 $\text{OP}_{\min, \text{stop}}$

We now consider the scenario where system (3.1) is controlled in a receding-horizon fashion by $\text{OP}_{\min, \text{stop}}$ as defined by Algorithm 2. That is, at each time instant $k \in \mathbb{Z}_{\geq 0}$, the first element of the optimal sequence $\mathbf{u}_{d(\phi_k)}^*(\phi_k)$ is calculated by $\text{OP}_{\min, \text{stop}}$, and then applied to system (3.1). This leads to the closed-loop system

$$x^+ \in f_{\mathcal{U}_\varepsilon^*(x)}(x) =: F_\varepsilon^*(x), \quad (3.17)$$

where $f_{\mathcal{U}_\varepsilon^*(x)}(x)$ is the set $\{f_u(x) : u \in \mathcal{U}_\varepsilon^*(x)\}$ and $\mathcal{U}_\varepsilon^*(x) := \{u_0 : \exists u_1, \dots, u_{d(x)} \in \mathcal{U} \text{ such that } V_\varepsilon(x) = J_{d(x)}(x, [u_0, \dots, u_{d(x)}])\}$ is the set of the first input of $d(x)$ -horizon

^{†1} This is a consequence of $\alpha_W(\sigma(x)) \leq \bar{\alpha}_V(\sigma(x))$, which holds from item (i) of Theorem 2.1, and the sublinear bounds $\bar{a}_W \mathbb{I} \leq \alpha_W$, $\bar{\alpha}_V \leq \bar{a}_V \mathbb{I} = (\bar{a}_V + \bar{a}_W) \mathbb{I}$. That is $\frac{\bar{a}_W}{\bar{a}_V + \bar{a}_W} \in (0, 1)$, hence $1 - \frac{\bar{a}_W}{\bar{a}_V + \bar{a}_W} \in (0, 1)$.

optimal input sequences at x , with $d(x)$ as defined in (3.6). We denote by $\phi(k, x)$ a solution to (3.17) at time $k \in \mathbb{Z}_{\geq 0}$ with initial condition $x \in \mathbb{R}^n$, with some abuse of notation.

We assume next that c_{stop} can be made small as desirable by taking $|\varepsilon|$ sufficiently small, which is with no loss of generality as we are free to design c_{stop} as wanted.

Assumption 3.1. There exists $\theta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$, with $\theta(\cdot, s) \in \mathcal{K}$ and $\theta(s, \cdot)$ non-decreasing for any $s > 0$, such that $c_{\text{stop}}(\varepsilon, x) \leq \theta(|\varepsilon|, \sigma(x))$ for any $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}^{n_\varepsilon}$. □

Example of functions c_{stop} which satisfy Assumption 3.1 are $c_{\text{stop}}(\varepsilon, x) = |\varepsilon|\sigma(x)$ and similar, like $c_{\text{stop}}(\varepsilon, x) = \max\{|\varepsilon_1|\alpha(\sigma(x)), |\varepsilon_2|\}$ for $\varepsilon = (\varepsilon_1, \varepsilon_2)$, $\alpha \in \mathcal{K}$ and $x \in \mathbb{R}^n$.

The next theorem provides stability guarantees for system (3.17).

Theorem 3.4. Consider system (3.17) and suppose c_{stop} verifies Assumption 3.1. There exists $\beta \in \mathcal{KL}$ such that, for any $\delta, \Delta > 0$, there exists $\varepsilon^* > 0$ such that for any $x \in \{z \in \mathbb{R}^n : \sigma(z) \leq \Delta\}$ and $\varepsilon \in \mathbb{R}^{n_\varepsilon}$ with $|\varepsilon| < \varepsilon^*$, solution $\phi(\cdot, x)$ to system (3.17) satisfies, for all $k \in \mathbb{Z}_{\geq 0}$

$$\sigma(\phi(k, x)) \leq \max\{\beta(\sigma(x), k), \delta\}. \quad (3.18)$$

□

Proof. The proof of Theorem 3.4 follows similar steps as the proofs of [24, Theorems 1-2] and Theorems 2.1-2.2. The difference is that the horizon in cost (3.7) is not fixed as in [24] or Chapter 2, and depends on the state. We circumvent this problem by showing that the infinite-horizon cost is a Lyapunov function as formalized next, which we believe is an interesting result on its own.

Proposition 3.5. Let $Y_\infty := V_\infty + W$, the following hold.

(i) For any $x \in \mathbb{R}^n$, $\underline{\alpha}_Y(\sigma(x)) \leq Y_\infty(x) \leq \bar{\alpha}_Y(\sigma(x))$, where $\underline{\alpha}_Y := \alpha_W$, $\bar{\alpha}_Y := \bar{\alpha}_V + \bar{\alpha}_W$, with $\alpha_W, \bar{\alpha}_V, \bar{\alpha}_W$ from SA2.

(ii) For any $x \in \mathbb{R}^n$, $\varepsilon \in \mathbb{R}^{n_\varepsilon}$, $v \in F_\varepsilon^*(x)$, $Y_\infty(v) - Y_\infty(x) \leq -\alpha_Y(\sigma(x)) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x))$ where $\alpha_Y = \alpha_W$, with α_W and $\bar{\alpha}_V$ from SA2, and c_{stop} comes from (3.6). □

Proof of Proposition 3.5. Let $\varepsilon \in \mathbb{R}^{n_\varepsilon}$, $x \in \mathbb{R}^n$ and $v \in F_\varepsilon^*(x)$, which is well-defined in view of Proposition 3.1. There exists $[u_0^*, u_1^*, \dots, u_{d(x)}^*] = \mathbf{u}_{d(x)}^*(x)$ such that $v = f_{u_0^*}(x)$ and $\mathbf{u}_{d(x)}^*(x)$ is an optimal input sequence for system (3.1) and cost (3.3) with horizon $d(x)$, which also verifies (3.6). Hence $V_\varepsilon(x) = J_{d(x)}(x, \mathbf{u}_{d(x)}^*(x))$.

From items (i) and (ii) of SA2, we have $Y_\infty(x) = V_\infty(x) + W(x) \leq \bar{\alpha}_V(\sigma(x)) + \bar{\alpha}_W(\sigma(x)) =: \bar{\alpha}_Y(\sigma(x))$. On the other hand, we have from item (ii) of SA2 that $\alpha_W(\sigma(x)) \leq W(x) + \ell_{u_0^*}(x)$ since $W(f_{u_0^*}(x)) \geq 0$. This implies that $\alpha_W(\sigma(x)) \leq W(x) + V_\varepsilon(x) \leq W(x) + V_\infty(x) = Y_\infty(x)$. Hence item (i) of Proposition 3.5 holds with $\underline{\alpha}_Y = \alpha_W$.

Consider the sequence $\hat{\mathbf{u}} := [u_1^*, u_2^*, \dots, u_{d(x)-1}^*, \bar{\mathbf{u}}]$ where $\bar{\mathbf{u}} := \mathbf{u}_\infty^*(\phi(d(x), x, \mathbf{u}_{d(x)}^*(x)|_{d(x)}))$, $\mathbf{u}_{d(x)}^*(x)|_{d(x)} = [u_0^*, \dots, u_{d(x)-1}^*]$ and ϕ denotes the solution of system (3.1). The sequence $\hat{\mathbf{u}}$ consists of the first $d(x)$ elements of $\mathbf{u}_{d(x)}^*(x)$ after u_0^* , followed by an optimal input sequence of infinite length at state $\phi(d(x), x, \mathbf{u}_{d(x)}^*(x)|_{d(x)})$, which exists according to SA1. Sequence $\bar{\mathbf{u}}$ minimizes $J_\infty(\phi(d(x), x, \mathbf{u}_{d(x)}^*(x)|_{d(x)}), \bar{\mathbf{u}})$ by virtue of SA1. From the definition of cost $J_{\gamma, d}$ in (3.3) and $V_\infty(v)$ in view of SA1,

$$V_\infty(v) \leq J_\infty(v, \hat{\mathbf{u}}) = J_{d(x)-1}(v, \hat{\mathbf{u}}|_{d(x)-1}) + J_\infty(\phi(d(x) - 1, v, \hat{\mathbf{u}}|_{d(x)-1}), \bar{\mathbf{u}}). \quad (3.19)$$

From Bellman optimality principle, we have $V_\varepsilon(x) = V_{d(x)}(x) = \ell_{u_0^*}(x) + V_{d(x)-1}(v) = \ell_{u_0^*}(x) + J_{d(x)-1}(v, \hat{\mathbf{u}}|_{d(x)-1})$, hence

$$J_{d(x)-1}(v, \hat{\mathbf{u}}|_{d(x)-1}) = V_\varepsilon(x) - \ell_{u_0^*}(x). \quad (3.20)$$

Moreover, by item (i) of SA2,

$$J_\infty(\phi(d(x) - 1, v, \hat{\mathbf{u}}|_{d(x)-1}), \bar{\mathbf{u}}) \leq \bar{\alpha}_V(\sigma(\phi(d(x) - 1, v, \hat{\mathbf{u}}|_{d(x)-1}))). \quad (3.21)$$

Consequently, in view of (3.19), (3.20) and (3.21),

$$V_\infty(v) \leq V_\varepsilon(x) - \ell_{u_0^*}(x) + \bar{\alpha}_V(\sigma(\phi(d(x) - 1, v, \hat{\mathbf{u}}|_{d(x)-1}))). \quad (3.22)$$

Since $\phi(d(x) - 1, v, \hat{\mathbf{u}}|_{d(x)-1}) = \phi(d(x), x, \mathbf{u}_{d(x)}^*(x)|_{d(x)})$ and (3.6) holds, $\sigma(\phi(d(x) - 1, v, \hat{\mathbf{u}}|_{d(x)-1})) = \sigma(\phi(d(x), x, \mathbf{u}_{d(x)}^*(x)|_{d(x)})) \leq c_{\text{stop}}(\varepsilon, x)$. Therefore,

$$V_\infty(v) \leq V_\varepsilon(x) - \ell_{u_0^*}(x) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x)). \quad (3.23)$$

By Theorem 2.3, $V_\varepsilon(x) \leq V_\infty(x)$, thus

$$V_\infty(v) \leq V_\infty(x) - \ell_{u_0^*}(x) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x)). \quad (3.24)$$

By invoking item (ii) of SA2, we derive $V_\infty(v) + W(v) \leq V_\infty(x) + W(x) - \alpha_W(\sigma(x)) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x))$, and since $Y_\infty = V_\infty + W$, the proof is completed with $\alpha_Y := \alpha_W$. ■

We now finalize the proof of Theorem 3.4. Let $\Delta, \delta > 0$. We select $\varepsilon^* > 0$ such that

$$\theta(\varepsilon^*, \underline{\alpha}_Y^{-1}(\tilde{\Delta})) < \bar{\alpha}_V^{-1}\left(\frac{1}{2}\tilde{\alpha}_Y(\tilde{\delta})\right), \quad (3.25)$$

where $\tilde{\alpha}_Y := \alpha_W \circ \bar{\alpha}_Y^{-1}$, $\tilde{\Delta} := \bar{\alpha}_Y(\Delta)$, $\tilde{\delta} := (\mathbb{I} - \frac{\tilde{\alpha}_Y}{2})^{-1} \circ \underline{\alpha}_Y(\delta)$ and θ comes from Assumption 3.1. Note that $(\mathbb{I} - \frac{\tilde{\alpha}_Y}{2})^{-1}$ is indeed of class \mathcal{K}_∞ as we assume without loss of generality that^{†2} $\mathbb{I} - \tilde{\alpha}_Y \in \mathcal{K}_\infty$, hence $\mathbb{I} - \tilde{\alpha}_Y + \frac{\tilde{\alpha}_Y}{2} \in \mathcal{K}_\infty$ and so is its inverse. Inequality (3.27) can always be verified by taking ε^* sufficiently small since $\theta(\cdot, \bar{\alpha}_V^{-1}(\tilde{\Delta})) \in \mathcal{K}$, and $\bar{\alpha}_V^{-1}(\frac{1}{2}\tilde{\alpha}_Y(\tilde{\delta})) > 0$. It follows from $\theta(\cdot, s) \in \mathcal{K}$ for any $s > 0$ and $\theta(s, \cdot)$ is non-decreasing for any $s \geq 0$, that $\theta(|\varepsilon|, \underline{\alpha}_Y^{-1}(s)) \leq \theta(\varepsilon^*, \underline{\alpha}_Y^{-1}(\tilde{\Delta}))$ for any $s \in [0, \tilde{\Delta}]$ and $|\varepsilon| < \varepsilon^*$. Furthermore, from Assumption 3.1 and item (i) of Proposition 3.5, we derive $c_{\text{stop}}(\varepsilon, x) \leq \theta(|\varepsilon|, \underline{\alpha}_Y^{-1}(Y_\infty(x)))$. Thus, in view of (3.25),

$$c_{\text{stop}}(\varepsilon, x) \leq \bar{\alpha}_V^{-1}\left(\frac{1}{2}\tilde{\alpha}_Y(\tilde{\delta})\right) \quad (3.26)$$

for any x such that $Y_\infty(x) \leq \tilde{\Delta}$. On the other hand, we have $\bar{\alpha}_V^{-1}(\frac{1}{2}\tilde{\alpha}_Y(\tilde{\delta})) \leq \bar{\alpha}_V^{-1}(\frac{1}{2}\tilde{\alpha}_Y(s))$ for any $s \in [\tilde{\delta}, \infty)$. Hence, for any $x \in \mathbb{R}^n$ such that $Y_\infty(x) \in [\tilde{\delta}, \tilde{\Delta}]$ and $|\varepsilon| < \varepsilon^*$,

$$\bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x)) \leq \frac{\tilde{\alpha}_Y(\tilde{\delta})}{2} \leq \frac{\tilde{\alpha}_Y(Y_\infty(x))}{2}. \quad (3.27)$$

Let $x \in \mathbb{R}^n$ with $\sigma(x) \leq \Delta$ and $v \in F_\varepsilon^*(x)$. In view of (3.26) and items (i) and (ii) of Proposition 3.5,

$$Y_\infty(v) - Y_\infty(x) \leq -\tilde{\alpha}_Y(Y_\infty(x)) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x)). \quad (3.28)$$

Since $\sigma(x) \leq \Delta$, $Y_\infty(x) \leq \bar{\alpha}_Y(\sigma(x)) \leq \bar{\alpha}_Y(\Delta) = \tilde{\Delta}$. Consider $Y_\infty(x) \in [0, \tilde{\delta})$. Since $c_{\text{stop}}(\varepsilon, x) \leq \bar{\alpha}_V^{-1}(\frac{1}{2}\tilde{\alpha}_Y(\tilde{\delta}))$ holds for $Y_\infty(x) \leq \tilde{\Delta}$, it holds here. Furthermore, since $\mathbb{I} - \tilde{\alpha}_Y \in$

^{†2} If that is not the case we can always find $\alpha' \in \mathcal{K}_\infty$ such that $\mathbb{I} - \tilde{\alpha}_Y \leq \mathbb{I} - \alpha'$. Indeed, as $\alpha_Y = \alpha_W = \underline{\alpha}_Y \leq \bar{\alpha}_Y$ holds from (i), which in turn implies $\alpha_Y \circ \bar{\alpha}_Y^{-1} \leq \mathbb{I}$, hence $s - \alpha_Y \circ \bar{\alpha}_Y^{-1}(s) \geq 0$ for all $s \geq 0$ and equality holds if and only if $s = 0$. Therefore there exists $\alpha' \in \mathcal{K}_\infty$ such that $\mathbb{I} - \alpha' \in \mathcal{K}_\infty$ and $\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1} \leq \mathbb{I} - \alpha'$, by Lemma B.1 [37]. A similar property is derived in [26].

\mathcal{K}_∞ holds without loss of generality, and in view of (3.28),

$$\begin{aligned} Y_\infty(v) &\leq Y_\infty(x) - \tilde{\alpha}_Y(Y_\infty(x)) + \bar{\alpha}_V(\theta(\varepsilon^*, \sigma(x))) \\ &\leq (\mathbb{I} - \tilde{\alpha}_Y)(\tilde{\delta}) + \frac{1}{2}\tilde{\alpha}_Y(\tilde{\delta}). \end{aligned}$$

Given the definition of $\tilde{\delta}$,

$$Y_\infty(v) \leq \left(\mathbb{I} - \frac{\tilde{\alpha}_Y}{2} \right) (\tilde{\delta}) = \underline{\alpha}_Y(\delta). \quad (3.29)$$

When^{†3} $Y_\infty(x) \geq \tilde{\delta}$, we derive from (3.27) that $-\tilde{\alpha}_Y(Y_\infty(x)) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x)) \leq -\frac{1}{2}\tilde{\alpha}_Y(Y_\infty(x))$. Thus, from (3.28),

$$Y_\infty(v) - Y_\infty(x) \leq -\frac{1}{2}\tilde{\alpha}_Y(Y_\infty(x)). \quad (3.30)$$

In view of (3.29) and (3.30), it follows for any $k \in \mathbb{Z}_{\geq 0}$ that

$$Y_\infty(\phi(k+1, x)) \leq \max \left\{ (\mathbb{I} - \frac{1}{2}\tilde{\alpha}_Y)(Y_\infty(x)), \underline{\alpha}_Y(\delta) \right\}, \quad (3.31)$$

where $\phi(k, x)$ is a solution starting at x for system (3.17). Furthermore, when $Y_\infty(x) \leq \underline{\alpha}_Y(\delta)$, $Y_\infty(v) \leq \underline{\alpha}_Y(\delta)$ follows. Indeed, if $Y_\infty(x) \in [\tilde{\delta}, \tilde{\Delta}]$, $Y_\infty(v) \leq Y_\infty(x) \leq \underline{\alpha}_Y(\delta)$ from (3.30), and if $Y_\infty(x) \in [0, \tilde{\delta})$, we deduce $Y_\infty(v) \leq \underline{\alpha}_Y(\delta)$ from (3.29). Hence the set $\{z \in \mathbb{R}^n : Y_\infty(z) \leq \underline{\alpha}_Y(\delta)\}$ is forward invariant for system (3.17). By iterating (3.31), we obtain

$$Y_\infty(\phi(k, x)) \leq \max \left\{ \tilde{\beta}(Y_\infty(x), k), \underline{\alpha}_Y(\delta) \right\}, \quad (3.32)$$

where $\tilde{\beta}(s, k) = (\mathbb{I} - \frac{1}{2}\tilde{\alpha}_Y)^{(k)}(s)$ for any $s \geq 0$, with $\tilde{\beta} \in \mathcal{KL}$ as $\lim_{k \rightarrow \infty} (\mathbb{I} - \frac{1}{2}\tilde{\alpha}_Y)^{(k)}(s) = 0$ for any $s \geq 0$ per the proof of item (iii) of Lemma B.1 from the Appendix, since $(\mathbb{I} - \frac{1}{2}\tilde{\alpha}_Y)(s) < s$ for $s > 0$ and $(\mathbb{I} - \frac{1}{2}\tilde{\alpha}_Y)(0) = 0$. Finally, invoking $\underline{\alpha}_Y(\sigma(x)) \leq Y_\infty(x) \leq \bar{\alpha}_Y(\sigma(x))$, we deduce

$$\sigma(\phi(k, x)) \leq \max \left\{ \underline{\alpha}_Y^{-1} \left(\tilde{\beta}(\bar{\alpha}_Y(\sigma(x)), k) \right), \delta \right\}. \quad (3.33)$$

Thus (3.18) holds with $\beta(s, k) = \underline{\alpha}_Y^{-1} \left(\tilde{\beta}(\bar{\alpha}_Y(s), k) \right)$ for any $s \geq 0$ and $k \in \mathbb{Z}_{\geq 0}$. ■

^{†3} It might be of interest to assume $\underline{\alpha}_Y(\delta) \leq \tilde{\Delta}$ as to the set $\{x : Y_\infty(x) \geq \tilde{\delta}\}$ be non-empty, but it is not necessary.

Theorem 3.4 provides a semiglobal practical stability property for set $\{z : \sigma(z) = 0\}$ for $\text{OP}_{\min, \text{stop}}$, like Theorem 3.3. In Theorem 3.3, we select budget B sufficiently large to provide a stability property, independent of the resulting $d(x)$ for $\text{OP}_{\min, \text{budget}}$. Instead, for $\text{OP}_{\min, \text{stop}}$, we select c_{stop} and its tuning parameter ε sufficiently small so that the algorithm stops whenever stability is guaranteed, without any further exploration.

Remark 3.7. In Theorem 3.3 we estimate a minimum horizon \bar{d} such that stability is preserved, and take a budget B sufficiently large to guarantee $d(x) \geq \bar{d}$. If we fix $c_{\text{stop}}(\varepsilon, x) = \bar{\alpha}_V^{-1} \circ \alpha_W (\mathbb{I} - \alpha_W \circ \bar{\alpha}_V^{-1})^{(\bar{d})}$, $\text{OP}_{\min, \text{stop}}$ will benefit of the same estimated stability guarantee. This follows despite $d(x) \leq \bar{d}$ for $\text{OP}_{\min, \text{stop}}$, see the proof of Proposition 3.1. Since the horizon for $\text{OP}_{\min, \text{stop}}$ is upper-bounded by \bar{d} , while \bar{d} is lower-bound for the horizon in $\text{OP}_{\min, \text{budget}}$, we can conclude that $\text{OP}_{\min, \text{stop}}$ requires less computational budget B than $\text{OP}_{\min, \text{budget}}$. \square

As for $\text{OP}_{\min, \text{budget}}$, by strengthening SA2, we can provide stronger properties. The following global exponential stability result is the equivalent of Corollary 3.1 derived for $\text{OP}_{\min, \text{budget}}$.

Corollary 3.2. *Suppose the following holds.*

(i) *There exist $\bar{a}_V, a_W > 0$, such that SA2 holds with $\bar{\alpha}_V \leq \bar{a}_V \cdot \mathbb{I}$, $\bar{\alpha}_W \leq \bar{a}_W \cdot \mathbb{I}$ and $\alpha_W \geq a_W \cdot \mathbb{I}$.*

(ii) *For any $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}^{n_\varepsilon}$, $c_{\text{stop}}(\varepsilon, x) \leq |\varepsilon| \sigma(x)$.*

Let ε^* such that

$$\varepsilon^* < \frac{a_W}{\bar{a}_V}. \quad (3.34)$$

Then, for any $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}^{n_\varepsilon}$ such that $|\varepsilon| \leq \varepsilon^*$, the solution $\phi(\cdot, x)$ to system (3.17) satisfies

$$\sigma(\phi(k, x)) \leq \frac{\bar{a}_V + \bar{a}_W}{a_W} \left(1 - \frac{a_W - |\varepsilon| \bar{a}_V}{\bar{a}_V + \bar{a}_W} \right)^k \sigma(x) \quad (3.35)$$

for all $k \in \mathbb{Z}_{\geq 0}$. \square

Proof. Let $x \in \mathbb{R}^n$. We select ε^* as in (3.34) and let $\varepsilon \in \mathbb{R}^{n_\varepsilon}$ such that $|\varepsilon| \leq \varepsilon^*$ and $v \in F_\varepsilon^*(x)$. In particular, we have shown that $Y_\infty(v) - Y_\infty(x) \leq -\frac{\mu}{\bar{a}_V + \bar{a}_W} Y_\infty(x)$ holds for

$\sigma(x) \in [0, L]$ in Corollary B.1. In this proof, we derive from (3.34) that $Y_\infty(v) - Y_\infty(x) \leq -\left(\frac{a_W - |\varepsilon|\bar{a}_V}{\bar{a}_V + \bar{a}_W}\right) Y_\infty(x)$ holds for any $\sigma(x) \geq 0$. Note that we do not require $\varepsilon^* < 1$ since $\bar{\alpha}_V(|\varepsilon|\sigma(x)) \leq \bar{a}_V|\varepsilon|\sigma(x)$ is guaranteed to hold for any $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}^{n_\varepsilon}$. We now proceed with the same argument as the proof of Corollary 2 in [55]. Let $x \in \mathbb{R}^n$ and denote $\phi(k, x)$ be a corresponding solution to (3.17) at time $k \in \mathbb{Z}_{\geq 0}$, it holds that $Y_\infty(\phi(k, x)) \leq \left(1 - \frac{a_W - |\varepsilon|\bar{a}_V}{\bar{a}_V + \bar{a}_W}\right)^k Y_\infty(x)$. Since $Y_\infty(x) \geq \alpha_W(\sigma(x)) \geq a_W\sigma(x)$ and $Y_\infty(x) \leq \bar{\alpha}_V(\sigma(x)) \leq (\bar{a}_V + \bar{a}_W)\sigma(x)$ holds from item (i) of Proposition 3.5 and item (i) of Corollary 3.2 for any $x \in \mathbb{R}^n$, it follows from $Y_\infty(\phi(k, x)) \leq \left(1 - \frac{a_W - |\varepsilon|\bar{a}_V}{\bar{a}_V + \bar{a}_W}\right)^k Y_\infty(x)$ that $a_W\sigma(\phi(k, x)) \leq \left(1 - \frac{a_W - |\varepsilon|\bar{a}_V}{\bar{a}_V + \bar{a}_W}\right)^k (\bar{a}_V + \bar{a}_W)\sigma(x)$ hence $\sigma(\phi(k, x)) \leq \frac{\bar{a}_V + \bar{a}_W}{a_W}\sigma(x) \left(1 - \frac{a_W - |\varepsilon|\bar{a}_V}{\bar{a}_V + \bar{a}_W}\right)^k$ and the proof is concluded. \blacksquare

Corollary 3.2 ensures a uniform global exponential stability property of set $\{x : \sigma(x) = 0\}$ for system (3.17). Indeed, $1 - \frac{a_W - |\varepsilon|\bar{a}_V}{\bar{a}_V + \bar{a}_W} \in (0, 1)$ in (3.35) as $|\varepsilon| < \frac{a_W}{\bar{a}_V}$ in view of (3.34), hence $\left(1 - \frac{a_W - |\varepsilon|\bar{a}_V}{\bar{a}_V + \bar{a}_W}\right)^k \rightarrow 0$ as $k \rightarrow \infty$. Furthermore, the estimated decay rate can be tuned via ε from 1 to $1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}$ as $|\varepsilon|$ decreases to zero.

Remark 3.8. Items (i)-(ii) of Corollary 3.2 are sufficient conditions for global exponential stability. If only global asymptotic stability is required, the stopping criterion can be selected as $c_{\text{stop}}(\varepsilon, x) \leq \bar{\alpha}_V^{-1}(\frac{1}{2}\alpha_W(\sigma(x)))$ for all $x \in \mathbb{R}^n$, where $\bar{\alpha}_V$ and α_W come from SA2. \square

Remark 3.9. As noted for $\text{OP}_{\text{min,budget}}$, it is possible to derive a semiglobal asymptotic stability property for $\text{OP}_{\text{min,stop}}$. We provide such result in Appendix B. When the sublinear properties in item (i) of Corollary B.1 are valid for $L = \infty$, we obtain Corollary 3.2. \square

By means of Corollary 3.2 and Corollary 3.1 we can better gauge the particularities of each algorithm. In one hand, we can chose $|\varepsilon|$ in Corollary 3.2 such that $1 - \frac{a_W - |\varepsilon|\bar{a}_V}{\bar{a}_V + \bar{a}_W} = e^{-\lambda}$ for the guaranteed^{†4} decay rate λ of Corollary 3.1. Hence, $\text{OP}_{\text{min,stop}}$ will provide the same guaranteed decay rate as $\text{OP}_{\text{min,budget}}$, however with a lower computational budget B as it will stop exploring as soon as the stopping criterion is verified. On the other hand, $\text{OP}_{\text{min,budget}}$ will exhaust the available budget and will “over-explore”. This might come

^{†4} It can be shown that λ lives in $(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}, 1)$

with a benefit in practice, we briefly come back to this point in an example, see Section 3.5.

3.4.3 Nominal robustness

As in Chapter 2.3.2, to ensure that the stability properties ensured so far are robust to so-called σ - or ρ -perturbations, as defined in e.g., [40], we can rely on two conditions according to [40, Theorem 2.8]. First, the set-valued mapping F_ε^* in (3.17), needs to be such that $F_\varepsilon^*(x)$ is nonempty and compact for any $x \in \mathbb{R}^n$. This is the case since \mathcal{U} and $d(x)$ are finite. Compactness of $F_\varepsilon^*(x)$ proceeds from the compactness of $\mathcal{U}_\varepsilon^*(x)$, $\mathcal{U}_\varepsilon^*(x)$ being a non-empty subset of finite set \mathcal{U} , when f is continuous as we assume in the upcoming lemma. This also applies for F_B^* in (3.13). Second, the Lyapunov function used to prove stability has to be continuous. In our case, the Lyapunov function constructed in Section 3.4.2, which applies for both $\text{OP}_{\min, \text{budget}}$ and $\text{OP}_{\min, \text{stop}}$, is $Y_\infty = V_\infty + W$. Since W is continuous by SA2, we need V_∞ to be continuous. The next proposition ensures this is the case under extra conditions on f , ℓ and σ . The result follows from [55, Theorem 3] with $\gamma = 1$ and $\mathcal{U} = \{1, \dots, M\}$, its proof is therefore omitted.

Lemma 3.1. *Suppose the following holds.*

- (i) *f and ℓ are continuous.*
- (ii) *For every $M \geq 0$, set $\{x : \sigma(x) \leq M\}$ is compact.*

Then V_∞ is continuous on \mathbb{R}^n . □

Item (ii) of Proposition 3.1 means that σ is radially unbounded, which is the case when it is given by, e.g. $\sigma = |\cdot|$, or more generally by $\sigma = |\cdot|_{\mathcal{A}}$ for a non-empty compact set \mathcal{A} .

3.4.4 Near-optimality guarantees

In Theorem 3.1 and 3.2, we have provided near-optimality guarantees of finite-horizon cost V_B and V_ε with respect to the infinite-horizon cost $V_\infty(x)$. This is an important feature of $\text{OP}_{\min, \text{budget}}$ and $\text{OP}_{\min, \text{stop}}$, but this does not directly provide us with information on the

actual value of the cost function (3.2) along solutions to (3.13) or (3.17). Indeed, in (3.13), we do not implement the whole sequence $\mathbf{u}_{d(x)}^*(x)$ given by $\text{OP}_{\min, \text{budget}}$ at x , instead we proceed in a receding horizon fashion. Similarly for (3.17) and $\text{OP}_{\min, \text{stop}}$. Therefore, we analyse a different cost.

3.4.4.1 $\text{OP}_{\min, \text{budget}}$

The relevant cost function to analyze is thus the running cost [25] defined as

$$\mathcal{V}_B^{\text{run}}(x) := \left\{ \sum_{k=0}^{\infty} \ell_{\mathcal{U}_B^*(\phi(k, x))}(\phi(k, x)) : \phi(\cdot, x) \text{ is a solution to (3.13)} \right\}, \quad (3.36)$$

where budget B guarantees $d(\phi(k, x)) > \bar{d}$ for all $k \in \mathbb{Z}_{\geq 0}$, and \bar{d} is a lower bound on the desired horizon at each step, which we can enforce by taking a sufficiently large budget according to Proposition 3.3. It has to be noted that $\mathcal{V}_B^{\text{run}}(x)$ is a set, since solutions of (3.13) are not necessarily unique. Each element $V_B^{\text{run}}(x) \in \mathcal{V}_B^{\text{run}}(x)$ corresponds then to the cost of a solution of (3.13). Clearly, $V_B^{\text{run}}(x)$ is not necessarily finite, as the stage costs may not decrease to 0 in view of Theorem 3.3. Indeed, only practical convergence is ensured in Theorem 3.3 in general. As a result, the corresponding running cost may not be finite. We therefore restrict our attention to the case where Corollary 3.1 holds, in the next theorem.

Theorem 3.5. *Consider system (3.13) and assume that Corollary 3.1 holds with tuple (λ, \bar{d}) . For any budget $B \geq \frac{M^{\bar{d}+2}-1}{M-1}$, $x \in \mathbb{R}^n$, and $V_B^{\text{run}}(x) \in \mathcal{V}_B^{\text{run}}(x)$,*

$$V_{\infty}(x) \leq V_B^{\text{run}}(x) \leq V_{\infty}(x) + w_B \cdot \sigma(x), \quad (3.37)$$

where $w_B := \left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right)^{\bar{d}} \left(\frac{\bar{a}_V + \bar{a}_W}{a_W}\right)^2 \frac{\bar{a}_V}{a_W(e^{\lambda} - 1)}$. □

Similarly to Theorem 3.1, the inequality $V_{\infty}(x) \leq V_B^{\text{run}}(x)$ of Theorem 3.5 directly follows from the optimality of $V_{\infty}(x)$. On the other hand, the inequality $V_B^{\text{run}}(x) \leq V_{\infty}(x) + w_B(x)$ provides a relationship between the running cost $V_B^{\text{run}}(x)$ and the infinite-horizon cost at state x , $V_{\infty}(x)$. The term w_B can be explicitly calculated, see Theorem 3.1 and [21, proof of Corollary 3] for the expressions of λ . The latter inequality in (3.37) confirms the intuition coming from Theorem 3.1 that a large computational budget B leads to tight near-optimality guarantees. That is, when $B \rightarrow \infty$, or equivalently $\bar{d} \rightarrow \infty$

according to Proposition 3.3, $w_B \rightarrow 0$ and $V_d^{\text{run}}(x) \rightarrow V_\infty(x)$, provided that budget B is chosen as to stabilize system (3.13). In contrast with Theorem 3.1, stability of system (3.13) plays a role in Theorem 3.5. Indeed, the term $\frac{1}{e^\lambda - 1}$ in (3.37) shows that the larger the exponential decay λ is, the smaller the error term w_B will be. This also applies for gain $\frac{\bar{a}_V + \bar{a}_W}{a_W}$ from (3.16) which appears squared in w_B . The running cost for the original OP was considered in [17], and it was found to perform at worst like the finite sequence, i.e. $V_B^{\text{run}}(x) \leq V_\infty(x) + \frac{\gamma^B}{1-\gamma}$. Compared to the bound derived for OP, the bound in Theorem 3.5 has similar benefits as Theorem 3.1, namely we are not limited to $\ell \in [0, 1]$, it does not explode to ∞ when $\gamma = 1$, and when $\sigma(x)$ is small follows $w_B \cdot \sigma(x)$ small. Moreover, the mismatch decays exponentially in B .

Remark 3.10. Inequality (3.37) can be written as a relationship of the finite-horizon costs in view of Theorem 3.5. In particular, we have $V_{d(x)}(x) \leq V_\infty(x) \leq V_B^{\text{run}}(x) \leq V_{d(x)}(x) + w_B \cdot \sigma(x)$, for any $x \in \mathbb{R}^n$. Hence, $V_{d(x)}(x)$ can be used to upper and lower bound $V_B^{\text{run}}(x)$ from the first call of $\text{OP}_{\text{min,budget}}$ at initial state x . \square

Other authors have considered the running cost of finite-horizon controllers applied in receding horizon fashion, like [25] in the context of Model Predictive Control (MPC). In particular, [25] derives relative performance of the running cost to the infinite-horizon optimal cost. From Theorem 3.5, we derive a similar result.

Corollary 3.3. *Suppose Theorem 3.5 holds for system (3.13) with tuple (λ, \bar{d}) . Then, for any $x \in \mathbb{R}^n$, budget $B \geq \frac{M^{\bar{d}+2}-1}{M-1}$, and $V_B^{\text{run}}(x) \in \mathcal{V}_B^{\text{run}}(x)$ such that $V_\infty(x) > 0$,*

$$\frac{V_B^{\text{run}}(x) - V_\infty(x)}{V_\infty(x) + W(x)} \leq \frac{w_B}{a_W}. \quad (3.38)$$

\square

Corollary 3.3 provides the relative performance of $V_B^{\text{run}}(x) + W(x)$ and $V_\infty(x) + W(x)$. When Assumption SA2 holds with $W \equiv 0$, Corollary 3.3 provides a relative relationship between any running cost V_B^{run} and the infinite-horizon cost V_∞ , as done in [25]. Interestingly, the obtained bound for relative performance is uniform in x . It does however conserve the desired properties of Theorem 3.5, in particular the exponential decay in \bar{d} in view of the expression of w_B in Theorem 3.5.

Remark 3.11. Corollary 3.3 follows from the stabilizability and detectability properties in Assumption 1, when the closed-loop system satisfies global exponential stability, while [25] results derive from relaxed dynamic programming, which relies on parameters of a modified Bellman optimality equation. Although a direct comparison between results is thus difficult, we note that our results are exponential in minimal horizon \bar{d} with rate $\left(1 - \frac{a_W}{\bar{a}_V + \bar{a}_W}\right)^{\bar{d}}$, while in [25] the obtained bound is of order $\frac{1}{(\eta+1)^{N-1}}$ for a parameter^{†5} η derived from the relaxed dynamic programming property and constant horizon $N + 1$. Hence, we expect that our bound may provide similar relative performance with a shorter horizon. We note that, while in [25] stability of the system is not assumed as in Corollary 3.3, a lower-bound for horizon $N + 1$ such that the results of [25] holds is expected, as in Corollary 3.3. \square

3.4.4.2 $\text{OP}_{\min, \text{stop}}$

As explained in the beginning of this section, the relationships between the finite-horizon cost V_ε and the infinite-horizon cost V_∞ provided in Section 3.4.1 are not sufficient. Indeed, we do not implement the whole sequence $\mathbf{u}_{d(x)}^*(x)$ given by OP_{\min} at x in (3.17), instead we proceed in a receding-horizon fashion. The relevant cost function to analyze in this context is the running cost [25] defined as

$$\mathcal{V}_\varepsilon^{\text{run}}(x) := \left\{ \sum_{k=0}^{\infty} \ell_{\mathbf{u}_\varepsilon^*(\phi(k,x))}(\phi(k,x)) : \phi(\cdot, x) \text{ is a solution to (3.17)} \right\}, \quad (3.39)$$

where $\ell_{\mathbf{u}_\varepsilon^*(\phi(k,x))}(\phi(k,x))$ is the actual stage cost incurred at time-step k . Each element $V_\varepsilon^{\text{run}}(x) \in \mathcal{V}_\varepsilon^{\text{run}}(x)$ corresponds then to the cost of a solution of (3.17). The same caveats for $\text{OP}_{\min, \text{budget}}$ applies here, that is, $\mathcal{V}_\varepsilon^{\text{run}}(x)$ is a set and is not necessarily finite, in view of Theorem 3.4. On the other hand, when the set $\{x \in \mathbb{R}^n : \sigma(x) = 0\}$ is globally exponentially stable as in Corollary 3.2, the elements of $\mathcal{V}_\varepsilon^{\text{run}}(x)$ in (3.39) are finite and satisfies the next property.

Theorem 3.6. Consider system (3.17) and suppose Corollary 3.2 holds. For any ε such that $|\varepsilon| < \varepsilon^*$, $x \in \mathbb{R}^n$, and $V_\varepsilon^{\text{run}}(x) \in \mathcal{V}_\varepsilon^{\text{run}}(x)$,

$$V_\infty(x) \leq V_\varepsilon^{\text{run}}(x) \leq V_\infty(x) + w_\varepsilon \sigma(x), \quad (3.40)$$

^{†5} We use η instead of γ , as used in [25], to not be confused with our use of γ as discount factor.

with $w_\varepsilon := \frac{\bar{a}_V}{a_W}(\bar{a}_V + \bar{a}_W)^2 \frac{|\varepsilon|}{a_W - \bar{a}_V|\varepsilon|}$, where constants come from Corollary 3.2. \square

Proof. Let $x \in \mathbb{R}^n$, $\varepsilon \in \mathbb{R}^{n\varepsilon}$ such that $|\varepsilon| < \varepsilon^*$ where ε^* is selected as in Corollary 3.2, $\phi(k+1, x) \in F_\varepsilon^*(\phi(k, x))$ for any $k \in \mathbb{Z}_{\geq 0}$ where ϕ is a solution to (3.17) initialized at x . The proof follows by following the steps of the proof of Theorem B.2, in particular inequality (B.10), however summed with $N \rightarrow \infty$. That is,

$$V_\varepsilon^{\text{run}}(x) \leq V_\varepsilon(\phi(0, x)) + \sum_{k=0}^{\infty} \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, \phi(k, x))). \quad (3.41)$$

All that remains is to compute a bound on $\sum_{k=0}^{\infty} \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, \phi(k, x)))$, which is possible by recalling that $\sigma(\phi(k, x)) \leq \frac{\bar{a}_V + \bar{a}_W}{a_W} \sigma(x) \left(1 - \frac{a_W - |\varepsilon| \bar{a}_V}{\bar{a}_V + \bar{a}_W}\right)^k$ holds from Corollary 3.2 and $\bar{\alpha}_V(c_{\text{stop}}(\varepsilon, \phi(k, x))) \leq \bar{a}_V |\varepsilon| \sigma(\phi(k, x))$ as the conditions of Corollary 3.2 are assumed to hold. Specifically, $\sum_{k=0}^{\infty} \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, \phi(k, x))) \leq |\varepsilon| \bar{a}_V \frac{(\bar{a}_V + \bar{a}_W)}{a_W} \sigma(x) \sum_{k=0}^{\infty} \left(1 - \frac{a_W - |\varepsilon| \bar{a}_V}{\bar{a}_V + \bar{a}_W}\right)^k$, which provides (3.40) as $\sum_{k=0}^{\infty} \left(1 - \frac{a_W - |\varepsilon| \bar{a}_V}{\bar{a}_V + \bar{a}_W}\right)^k = \frac{\bar{a}_V + \bar{a}_W}{a_W - \bar{a}_V |\varepsilon|}$. The lower bound $V_\infty(x) \leq V_\varepsilon^{\text{run}}(x)$ follows from the optimality of $V_\infty(x)$. Since (3.41) holds for an arbitrary solution of (3.17), $\phi(k+1, x) = f_{u_k^r}(\phi(k, x))$ for any $k \in \mathbb{Z}_{\geq 0}$, the resulting bound holds for any $V_\varepsilon^{\text{run}}(x) \in \mathcal{V}_\varepsilon^{\text{run}}(x)$. \blacksquare

The inequality $V_\infty(x) \leq V_\varepsilon^{\text{run}}(x)$ of Theorem 3.6 directly follows from the optimality of V_∞ . The inequality $V_\varepsilon^{\text{run}}(x) \leq V_\infty(x) + w_\varepsilon(x)$ provides a relationship between the running cost $V_\varepsilon^{\text{run}}(x)$ and the infinite-horizon cost at state x , $V_\infty(x)$. The latter inequality in (3.40) confirms the intuition coming from Theorem 2.3 that a smaller stopping criterion leads to tighter near-optimality guarantees. That is, when $|\varepsilon| \rightarrow 0$, $w_\varepsilon \rightarrow 0$ and $V_\varepsilon^{\text{run}}(x) \rightarrow V_\infty(x)$ for any $x \in \mathbb{R}^n$, provided that Corollary 3.2 holds. In contrast with Theorem 2.3, stability of system (3.17) plays a role in Theorem 3.6. Indeed, the term $\frac{1}{a_W - \bar{a}_V |\varepsilon|}$ in the expression of w_ε shows that the running cost is large when $|\varepsilon|$ is close to $\frac{a_W}{\bar{a}_V}$, hence when stability is not guaranteed the running cost might explode.

One interesting distinction between Theorem 3.6 and Theorem 3.5 is that error term w_B decays exponentially in \bar{d} , the minimum possible horizon under budget B , while w_ε decays “only” linearly in $|\varepsilon|$. This however does not imply that $\text{OP}_{\text{min}, \text{budget}}$ provides better running cost guarantees, as: (i) the budget potentially increases exponentially in \bar{d} ; (ii) we

can fit c_{stop} or $|\varepsilon|$ such that $w_\varepsilon \leq w_B$. Indeed, the guaranteed bound in Theorem 3.6 is less conservative due to the stopping criterion. However, the fact that $\text{OP}_{\text{min,budget}}$ might explore larger horizons, as it always exhaust budget B to maximize $d(x)$, might provide a better observed running cost than $\text{OP}_{\text{min,stop}}$ with an ill fitted c_{stop} .

Remark 3.12. The running cost for the original OP was considered in [17], and it was found to perform at worst like the finite sequence, i.e. $V_{\gamma,\bar{d}}^{\text{run}}(x) \leq V_{\gamma,\infty}(x) + \frac{\gamma^{\bar{d}}}{1-\gamma}$, where OP calculates at x an input sequence $\mathbf{u}_{\gamma,\bar{d}}^*(x)$ with *discounted* cost $V_{\gamma,\bar{d}}(x)$ for some horizon $\bar{d} \in \mathbb{Z}_{>0}$ and discount factor $\gamma \in (0, 1)$. Compared to the bound derived for OP, the bound in Theorem 3.6 has similar benefits as Theorem 3.5, namely: we are not limited to $\ell \in [0, 1]$, it is finite for undiscounted costs ($\gamma = 1$), and when $\sigma(x)$ is small follows $w_\varepsilon \cdot \sigma(x)$ small. Moreover, the mismatch is smaller for faster decays, i.e. for smaller $1 - \frac{a_W - |\varepsilon|\bar{a}_V}{\bar{a}_V + \bar{a}_W}$. \square

Remark 3.13. Inequality (3.40) can be written as a relationship of the finite-horizon costs in view of Theorem 3.6. In particular, we have $V_\varepsilon(x) \leq V_\infty(x) \leq V_\varepsilon^{\text{run}}(x) \leq V_\varepsilon(x) + w_\varepsilon \cdot \sigma(x)$, for any $x \in \mathbb{R}^n$. Hence, $V_\varepsilon(x)$, which is returned by the algorithm at the initial time, can be used to upper and lower bound $V_\varepsilon^{\text{run}}(x)$ from the first call of OP_{min} at initial state x . \square

Remark 3.14. In Appendix B.3 we define and analyse two different types of costs, namely the average running cost and the “cost-to-attractor”, which are defined even when only practical stability is upheld. \square

From Theorem 3.6, we derive a relative performance of the running cost to the infinite-horizon optimal cost, similar to [25].

Corollary 3.4. *Suppose the conditions of Theorem 3.6 hold for system (3.17). Then, for any $x \in \mathbb{R}^n$ such that $V_\infty(x) > 0$, $\varepsilon \in \mathbb{R}^{n_\varepsilon}$ such that $|\varepsilon| < \varepsilon^*$, and $V_\varepsilon^{\text{run}}(x) \in \mathcal{V}_\varepsilon^{\text{run}}(x)$,*

$$\frac{V_\varepsilon^{\text{run}}(x) - V_\infty(x)}{V_\infty(x) + W(x)} \leq \frac{w_\varepsilon}{a_W}, \quad (3.42)$$

where w_ε and a_W come from Theorem 3.6. \square

Proof. Let $x \in \mathbb{R}^n$, $\varepsilon \in \mathbb{R}^{n_\varepsilon}$ such that $|\varepsilon| < \varepsilon^*$ where ε^* is selected as in (3.34). Since Theorem 3.6 holds, $V_\varepsilon^{\text{run}}(x) \leq V_\infty(x) + w_\varepsilon \cdot \sigma(x)$. Moreover, $V_\varepsilon(x) \leq V_\infty(x)$ holds from Theorem 3.1. Thus by invoking upper-bound $a_W \sigma(x) \leq V_\infty(x) + W(x)$ from item (i) of

Proposition 3.5, it follows that $\sigma(x) \leq \frac{V_\infty(x)+W(x)}{a_W}$ hence $V_\varepsilon^{\text{run}}(x) \leq V_\varepsilon(x) + (V_\infty(x) + W(x))\frac{w_\varepsilon}{a_W}$, which implies (3.42). \blacksquare

Corollary 3.4 provides a bound on the relative performance of $V_\varepsilon^{\text{run}}(x) + W(x)$ and $V_\infty(x) + W(x)$. When SA2 holds with $W \equiv 0$, Corollary 3.4 gives a relative relationship between any running cost $V_\varepsilon^{\text{run}}$ and the infinite-horizon cost V_∞ , as done in [25]. Interestingly, the obtained bound is uniform in x . It does however conserve the desired properties of Theorem 3.6, in particular that it can be made as small as desired by tuning ε , in view of the expression of w_ε in Theorem 3.6.

Remark 3.15. Corollary 3.4 follows from the stabilizability and detectability properties in SA2, when the closed-loop system satisfies global exponential stability, while [25] results derive from relaxed dynamic programming, which relies on parameters of a modified Bellman optimality equation. Moreover, we only implicitly rely on horizon $d(x)$ to provide an estimate on the running cost, and our main parameter is given by ε to control c_{stop} . Hence an indirect comparison to [25] like done in Remark 3.11 is not possible.

We note however that our results are linear in $|\varepsilon|$ and decrease with faster decay rate $1 - \frac{a_W - |\varepsilon|a_V}{a_V + a_W}$, while in [25] the obtained bound is of order $\frac{1}{(\xi+1)^{N-1}}$ for a parameter^{†6} ξ derived from the relaxed dynamic programming property and constant horizon $N + 1$. We also note that, while in [25] stability of the system is not assumed as in Corollary 3.4, a necessary lower-bound for horizon $N + 1$ such that the results of [25] hold is expected, as in Corollary 3.4. \square

In the following section, we will illustrate by means of an example our previous results on the running cost for both flavors of OP_{\min} .

^{†6} We use ξ instead of γ , as used in [25], to not be confused with a discount factor γ .

3.5 Example: a cubic integrator

System. We consider the cubic integrator from [24, Example 1], i.e.

$$\begin{aligned}x_1^+ &= x_1 + u \\x_2^+ &= x_2 + u^3\end{aligned}$$

where $(x_1, x_2) := x \in \mathbb{R}^2$ and $u \in \mathbb{R}$. It was verified in [24] that an open-loop sequence of inputs drives the system to $x = 0$ in a finite number of steps. This open-loop sequence can be expressed as three feedback gains $K_1(x) = -x_1$, $K_2(x) = x_2^{\frac{1}{3}}$ and $K_3(x) = \left(-\frac{1}{2} + \sqrt{\frac{7}{12}}\right) x_2^{\frac{1}{3}}$, which are successively applied.

We propose here to switch between these gains to minimize cost (3.2), with $\ell_u(x) = |x_1|^3 + |x_2| + |K_u(x)|^3$ for any $x \in \mathbb{R}^2$ and $u \in \{1, 2, 3\}$. Note that we cannot design a local LQR controller for this system, due to the lack of stabilizability of the linearized model at the origin. We therefore consider the switched system

$$\begin{aligned}x_1^+ &= x_1 + K_u(x) \\x_2^+ &= x_2 + (K_u(x))^3\end{aligned}\tag{3.43}$$

with $u \in \{1, 2, 3\}$.

OP_{min,budget}. We first apply OP_{min,budget} to illustrate the near-optimality and stability guarantees of Section 3.4. To do this, note that SA1 applies for the same reasons as in [24]. By taking $\sigma(x) = |x_1|^3 + |x_2|$ for any $x \in \mathbb{R}^2$, Assumption SA2 holds $\alpha_W = \mathbb{I}$, $W = \bar{\alpha}_W = 0$ and $\bar{\alpha}_V = 14\mathbb{I}$, as in [24]. We verify Corollary 3.1 conditions with $a_W = 1$, $\bar{a}_V = 14$ and $\bar{a}_W = 0$, and conclude that for any budget $B \geq \frac{3^{73}-1}{2}$ ensures global exponential stability. Consequently, Theorem 3.5 also holds. The lower-bound on B is conservative, as the horizon \bar{d} itself in Corollary 3.1 is subject to some conservatism, and that OP_{min,budget} will ensure large horizons for smaller budgets in general.

We have thus fixed the budget to $B = 3000$ for initial condition $x = (-1, 1.5)$. Figure 3.1 shows the evolution of the state, and we see that both x_1 and x_2 converge to zero, as ensured by Corollary 3.1. We then consider several initial conditions and we study the impact of the budget on the actual running cost estimated by running simulations over

200 steps. We see in Table 3.1 that the estimated running cost becomes smaller when increasing the budget, which is consistent with Theorem 3.5. In other words, the larger the budget, the better the running cost performance.

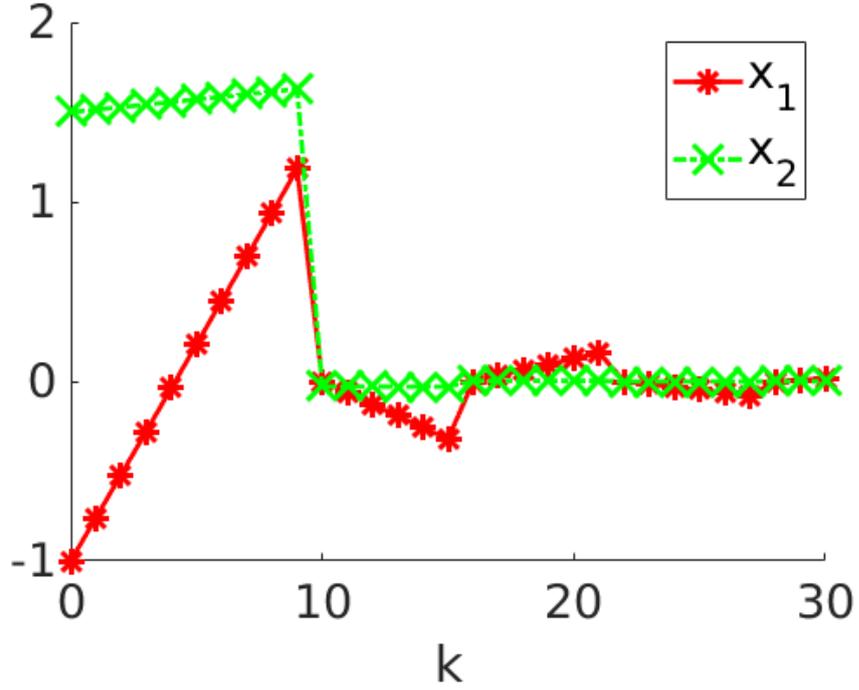


Figure 3.1: State evolution for $OP_{\min,budget}$ with $B = 3000$ and $x = (-1, 1.5)$.

	Budget		
	30	300	3000
(10, 15)	199015	13757	12609
Initial (-1 , 1.5)	314	28	22
States (-15, -10)	128184477	46875	42952
(10, -15)	14180	2802	2615

Table 3.1: Estimated running cost of $OP_{\min,budget}$ for various budgets and initial conditions.

OP_{min,stop}. We also run an experiment with OP_{min,stop}. By taking $c_{\text{stop}}(\varepsilon, x) = \varepsilon\sigma(x)$ with $\varepsilon \in \mathbb{R}_{\geq 0}$, we verify the conditions of Corollary 3.2 and Theorem 3.6, hence OP_{min,stop} can provide stability and running cost guarantees. In particular, by taking $\varepsilon = \frac{aw}{2\bar{a}_V} = \frac{1}{28} = 0.035$, we estimate that $\sigma(\phi(k, x)) \leq 14 \left(\frac{27}{28}\right)^k \sigma(x)$ and that $V_\infty(x) - V_\varepsilon^{\text{run}}(x) \leq w_\varepsilon\sigma(x)$ holds with $w_\varepsilon = \bar{a}_V^2 = 196$. For such ε and initial state $x = (-1, 1.5)$, we observe the same response as OP_{min,budget} in Figure 3.1.

Similarly to OP_{min,budget}, these estimates are subject to conservatism, and in fact OP_{min,stop} finds the best input sequence observed with ε as high as $\frac{1}{9}$, and convergence to the origin for ε as high as $\frac{10}{12} = 0.833$. However, this conservatism is objectively less of an issue compared to the ones observed for OP_{min,budget}. First, the bound $B \geq \frac{3^{73}-1}{2}$ is too separated to the empirically estimated budget $B = 3000$, while bound $\varepsilon < \frac{1}{14} = 0.071$ is about the same order of magnitude to the empirically estimated parameter $\varepsilon = \frac{1}{9} = 0.111$. Moreover, the chosen parameter $\varepsilon = 0.035$ can be put to practice for OP_{min,stop}, as the required budget is in average $B_{\text{avg}} = 4336$, with a maximum budget $B_{\text{max}} = 8506$. OP_{min,budget} is uncomputable for $B = \frac{3^{73}-1}{2}$.

In Table 3.2, we provide the estimates of $V_\varepsilon^{\text{run}}(x)$ for $x = (10, 15)$ for different choices of ε . We also provide the associated computational budget utilized by OP_{min,stop}. The calculated running cost becomes smaller when we decrease parameter ε , which is consistent with Theorem 3.6. We also see how the computational budget B adapts to fulfill the stopping criterion. In Appendix C, we provide the state plot for some of these ε , to better illustrate the different sequences found by OP_{min,stop}.

Remark 3.16. Example 1 helps to illustrate the main differences between OP_{min,budget} and OP_{min,stop}. First, the execution time for OP_{min,budget} is deterministic since the budget is fixed, hence suitable for real-time application constraints. However, without a more precise knowledge of the branching factor, denoted as κ in [32], we are overly conservative in our estimate of computational budget B . Second, OP_{min,stop} side-steps the issue and stops as soon as a “good” sequence is found. However, this implies that the execution time of OP_{min,stop} might expand on “harder” states. Fortunately, in light of the analysis of this chapter, we can show that OP_{min,stop} will at worst only require the same budget as OP_{min,budget}, e.g. take c_{stop} that guarantees OP_{min,stop} stops at horizon \bar{d} . \square

ε	$V_\varepsilon^{\text{run}}(x)$	B_{max}	B_{avg}
5	∞	2	2
1.111	199020*	26	2.1
0.909	57770	460	159.4
0.833	22697	460	137.7
0.588	13757	790	222.2
0.142	13757	4762	2289.5
0.111	12609	4762	1986.6
0.071	12609	4762	1986.6
0.035	12609	9095	4503.9
0.001	12609	14695	11265

Table 3.2: Estimated running cost of $\text{OP}_{\text{min,stop}}$ and associated computational budget utilized by $\text{OP}_{\text{min,stop}}$ for various ε and initial state $x = (10, 15)$. Symbol ' ∞ ' signifies that the state diverges to infinity, and symbol '*' is there to signalize that the state fails to converge but does not diverge to infinity.

Remark 3.17. In this example, $\text{OP}_{\text{min,stop}}$ with $\varepsilon = \frac{1}{9}$ sometimes require a budget of 4762, which is higher to $\text{OP}_{\text{min,budget}}$ with budget $B = 3000$. In this case, we can think that $\text{OP}_{\text{min,budget}}$ is “under spending”. On the other hand, $\text{OP}_{\text{min,budget}}$ is “over spending” 1013 budget in average. Interestingly, they both attain the same input sequence, hence same performance.

It is unclear if this over spending of budget in average can balance out a sporadic under spending of budget. That is, in some sense, if solving a harder optimal control problem, at average, preserves stability in case the required optimal problem is not fully solved, at given times. From another point of view, one might be interested in the scenario whether the full budget (or the stopping criterion) cannot be fulfilled due to sporadic lack of computational power, e.g. the computational power has to be used elsewhere, like treat a higher priority event. This is in our scope for future work. \square

3.6 Case study

3.6.1 Main result

We apply the previous results to the scenario where we are given a finite number of feedback laws, among which one is optimal in a neighborhood of the origin and one ensures a global exponential stability property. Our objective is to optimally switch among these controllers to minimize a given cost. This problem formulation is inspired by works on uniting controllers e.g., [57–59] where the total of number of controllers is two (one locally optimal and one global) and the switching occurs when the state is close to the attractor. Here, we allow for a larger number of controllers, and, more importantly, we exploit $\text{OP}_{\min, \text{stop}}$ to select a given feedback law at any time instant along solutions.

We concentrate on the next class of systems

$$x^+ = f(x, \kappa) \tag{3.44}$$

$$\kappa = g(u, x), \tag{3.45}$$

where $x \in \mathbb{R}^{n_x}$ is the state, $\kappa \in \mathbb{R}^{n_\kappa}$ is the feedback law output, which is parameterized by $u \in \{1, \dots, M\}$ with $M \geq 2$, $n_\kappa \in \mathbb{Z}_{>0}$. Vector field f is assumed to be continuously differentiable. We focus on quadratic infinite-horizon costs of the form

$$J_\infty(x, \mathbf{u}) = \sum_{k=0}^{\infty} x_k^\top Q x_k + \kappa_k^\top R \kappa_k, \tag{3.46}$$

where $k \in \mathbb{Z}_{\geq 0}$ is the time-step, $Q \in \mathbb{R}^{n_x \times n_x}$ and $R \in \mathbb{R}^{n_\kappa \times n_\kappa}$ are symmetric and positive definite matrices, and $\mathbf{u} := [u_0, u_1, \dots]$ is a sequence of controller choices. We aim to minimize (3.46) over the choices of u . We assume that set $\{1, \dots, M\}$ contains at least one controller, corresponding to index u_{global} , which globally exponentially stabilizes the origin of the induced closed-loop system (3.44) in the following sense.

Assumption 3.2. There exists $u_{\text{global}} \in \{1, \dots, M\}$, $P \in \mathbb{R}^{n_x \times n_x}$ symmetric, positive definite matrix and $a, b > 0$ such that, for $V : x \mapsto x^\top P x$, the following hold for any $x \in \mathbb{R}^n$.

(i) $V(f(x, g(u_{\text{global}}, x))) - V(x) \leq -aV(x)$.

(ii) $|g(u_{\text{global}}, x)| \leq b|x|$. □

Item (i) of Assumption 3.2 implies that V is an exponential Lyapunov function for system $x^+ = f(x, g(u_{\text{global}}, x))$. Item (ii) of Assumption 3.2 means that the norm of the feedback law is upper-bounded by a linear term in $|x|$, which is the case when $g(u_{\text{global}}, x)$ is linear for instance. Design techniques to verify Assumption 3.2 can be found in, e.g., [4, 18, 20, 34], for given classes of systems.

To design the local optimal feedback law, we rely on the next assumption.

Assumption 3.3. Let $A := \frac{\partial f}{\partial x}|_{(0,0)}$ and $B := \frac{\partial f}{\partial u}|_{(0,0)}$. The pair (A, B) is stabilizable. \square

In view of Assumption 3.3, we can design the optimal controller for the linearized model of (3.44) at the origin, i.e. $x^+ = Ax + Bu$ and cost $\ell(x, \kappa) = x^\top Qx + \kappa^\top R\kappa$, where Q, R come from (3.46). This local controller is given by $g(u_{\text{local}}, x) = -Kx$ for any $x \in \mathbb{R}^m$ with $K := (B^\top P_{\text{local}}B + R)^{-1}B^\top P_{\text{local}}A$. Matrix P_{local} is the unique solution of the discrete Riccati equation $P_{\text{local}} = A^\top P_{\text{local}}A - A^\top P_{\text{local}}B(R + B^\top P_{\text{local}}B)^{-1}B^\top P_{\text{local}}A + Q$.

Remark 3.18. The conditions of Assumption 3.2 can be relaxed. Possible extensions include the cases where the stage cost in (3.46) is only quadratic in a neighborhood of the origin, or where the global controller only ensures stability properties with respect to a neighborhood of the origin. These extensions require more investigations and are left for future work. \square

We choose the stopping criterion as $c_{\text{stop}}(\varepsilon, x) = \varepsilon x^\top Px$, with P from Assumption 3.2, for any $x \in \mathbb{R}^{n_x}$ and some $\varepsilon \in \mathbb{R}_{>0}$.

The next statement guarantees that the standing assumptions stated in Section 3.3.2 are verified.

Proposition 3.6. Consider system (3.44) where Assumptions 3.2 is verified. The following hold.

(i) SA1 is verified.

(ii) SA2 is verified with $\sigma(x) = x^\top Px$ for any $x \in \mathbb{R}^n$, $\alpha_W := \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}$, $\bar{\alpha}_V := \frac{\nu_1}{1-e^{-\nu_2}}\mathbb{I}$ for $\nu_1 := \frac{\lambda_{\max}(Q)+b\lambda_{\max}(R)}{\lambda_{\min}(P)}$, $\nu_2 := \ln(1-a)^{-1}$ and $W = \bar{\alpha}_W := 0$.

(iii) SA3 and Assumption 3.1 are verified with $\theta(\varepsilon, \sigma(x)) = \varepsilon\sigma(x)$, for any $x \in \mathbb{R}^n$. \square

Proof. Let $x \in \mathbb{R}^n$. From item (i) of Assumption 3.2, there exists $P \in \mathbb{R}^n \times \mathbb{R}^n$ symmetric, positive definite matrix and $a > 0$ such that, for solution $\phi_{\text{global}}(k, x)$ to system (3.44) with feedback law u_{global} initialized at x , the following holds for $k \in \mathbb{Z}_{\geq 0}$.

$$\phi_{\text{global}}(k+1, x)^\top P \phi_{\text{global}}(k+1, x) - \phi_{\text{global}}(k, x)^\top P \phi_{\text{global}}(k, x) \leq -a \phi_{\text{global}}(k, x)^\top P \phi_{\text{global}}(k, x). \quad (3.47)$$

Hence, $\phi_{\text{global}}(k+1, x)^\top P \phi_{\text{global}}(k+1, x) \leq (1-a) \phi_{\text{global}}(k, x)^\top P \phi_{\text{global}}(k, x)$. By iteration and recalling that $\phi_{\text{global}}(0, x) = x$, we derive

$$\sigma(\phi_{\text{global}}(k+1, x)) \leq (1-a)^k \sigma(x), \quad (3.48)$$

where $\sigma(x) = x^\top P x$. We show next that $\ell_{u_{\text{global}}}(\phi_{\text{global}}(k, x)) \leq \nu_1 \sigma(x) e^{-\nu_2 k}$ for some $\nu_1, \nu_2 > 0$ and $\ell_{u_{\text{global}}}(x) := x^\top Q x + g(u_{\text{global}}, x)^\top R g(u_{\text{global}}, x)$. Since Q is positive definite, $x^\top Q x \leq \lambda_{\max}(Q) |x|^2$. Furthermore $|x|^2 \leq \frac{1}{\lambda_{\min}(P)} x^\top P x$, hence $x^\top Q x \leq \frac{\lambda_{\max}(Q)}{\lambda_{\min}(P)} x^\top P x$. Similarly for R and invoking item (ii) of Assumption 3.2, we obtain that, $g(u_{\text{global}}, x)^\top R g(u_{\text{global}}, x) \leq \lambda_{\max}(R) |g(u_{\text{global}}, x)|^2 \leq \lambda_{\max}(R) b |x|^2 \leq \frac{b \lambda_{\max}(R)}{\lambda_{\min}(P)} x^\top P x$.

It follows from (3.48) that

$$\ell_{u_{\text{global}}}(\phi_{\text{global}}(k, x)) \leq \nu_1 \sigma(x) e^{-\nu_2 k}, \quad (3.49)$$

where $\nu_1 = \frac{\lambda_{\max}(Q) + b \lambda_{\max}(R)}{\lambda_{\min}(P)}$ and $\nu_2 = \ln(1-a)^{-1}$. By invoking [24, Lemma 1], we derive that item (i) of SA2 holds with $\bar{\alpha}_V := \frac{\nu_1}{1-e^{-\nu_2}} \mathbb{I}$. In particular, $V_\infty(x) \leq \bar{\alpha}_V(\sigma(x)) = \frac{\nu_1}{1-e^{-\nu_2}} x^\top P x$, which is finite for all $x \in \mathbb{R}^n$, thus [39, Theorem 2] is verified and SA1 holds, hence item (i) of Proposition 3.6 is verified. On the other hand, since $\lambda_{\min}(Q) |x|^2 \leq x^\top Q x$ holds, $\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} \sigma(x) = \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} x^\top P x \leq x^\top Q x \leq \ell_u(x)$ for any $u \in \{1, \dots, M\}$, hence item (ii) of SA2 is verified with $\alpha_W := \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} \mathbb{I}$ and $W = \bar{\alpha}_W = 0$. We have proved that item (ii) of Proposition 3.6 holds. Item (iii) of Proposition 3.6 follows immediately since $c_{\text{stop}}(\varepsilon, x) = \theta(|\varepsilon, \sigma(x)) := \varepsilon \sigma(x)$ by our choice of c_{stop} , hence Assumption 3.1 holds. Furthermore, $\varepsilon \sigma(x) > 0$ when $\sigma(x) > 0$ follows from $\varepsilon \in \mathbb{R}_{>0}$, hence SA3 holds. \blacksquare

As a result, we can tune ε according to Corollary 3.2 to endow the corresponding system (3.44) with global exponential stability and performance guarantees as formalized next.

Proposition 3.7. Consider system (3.44) where Assumptions 3.2 is verified. Let $\varepsilon \in (0, \frac{a_W}{\bar{a}_V})$, where $\bar{a}_V = \frac{\nu_1}{1-e^{-\nu_2}}$, $a_W = \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}$ with Q from (3.46), P from Assumption 3.2 and ν_2, ν_1 from Proposition 3.6. Let $\sigma(x) = x^\top P x$ for any $x \in \mathbb{R}^{n_x}$. The following hold for any $x \in \mathbb{R}^n$.

$$(i) \text{ For any } k \in \mathbb{Z}_{\geq 0}, \sigma(\phi(k, x)) \leq \frac{\bar{a}_V}{a_W} \left(1 + \varepsilon - \frac{a_W}{\bar{a}_V}\right)^k \sigma(x).$$

$$(ii) |V_\varepsilon^{\text{run}}(x) - V_\infty(x)| \leq \frac{\bar{a}_V^3}{a_W} \frac{\varepsilon}{a_W - \bar{a}_V \varepsilon}. \quad \square$$

Proposition 3.7 is the application of Corollary 3.2 and Theorem 3.6, its proof is therefore omitted.

Remark 3.19. Note that we do not show that we recover the properties of the local optimal controller in a neighborhood of the origin. This is work in progress, however we show in simulations that this is indeed the case. \square

3.6.2 Example: a flexible joint robotic arm

We consider the flexible joint robotic arm model from [77, Section 4], discretized by an Euler scheme with $T > 0$, that gives

$$x^+ = x + T(A_p x + B_p \kappa - E_p \phi(x)), \quad (3.50)$$

where $x \in \mathbb{R}^4$, $\kappa \in \mathbb{R}^1$,

$$A_p = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -48.6 & -1.25 & 48.6 & 0 \\ 0 & 0 & 0 & 1 \\ 19.5 & 0 & -16.17 & 0 \end{bmatrix}, \quad (3.51)$$

$B_p = (0, 21.6, 0, 0)$, $E_p = (0, 0, 0, 3.33)$. System (3.50) has a nonlinearity of type Lur'e, due term $\phi(x) = x_3 + \sin(x_3)$. We fix $T = 0.1$. The infinite-horizon cost is given by

$$J_\infty(x, \kappa) := \sum_{k=0}^{\infty} x_k^\top Q x_k + R \kappa_k^2, \quad (3.52)$$

where $Q = I_{4 \times 4}$, with $I_{4 \times 4}$ is the identity matrix of dimension 4, and $R = 1$.

The jump map of system (3.50) is continuously differentiable at the origin, and the linearized system at the origin is controllable. Hence, Assumption 3.3 is verified and we design a local controller $g(1, x) = -K_{\text{local}}x$ that optimizes (3.52) for linear system of (3.50) around the origin ($u_{\text{local}} = 1$), as in Section 3.6.A. On the other hand, we design the global controller $g(2, x) = -K_{\text{global}}x$ with $K_{\text{global}} = [3.6, 0.9, -1.5, 0.3]$, which verifies LMI conditions found in [20, Theorem 2], and $u_{\text{global}} = 2$. In fact, we slightly modify the LMI^{†7} in [20, Theorem 2] to guarantee $V(f(x, -K_{\text{global}}x)) - V(x) \leq -a'(x^\top Qx + x^\top K_{\text{global}}RK_{\text{global}}x)$ for some $a' > 0$. This is done to ensure a less conservative estimate than Proposition 3.7 and conclude that cost (3.46) for u_{global} is given (and upper-bounded) by $\sum_{k=0}^{\infty} x_k^\top Qx_k + x_k^\top K_{\text{global}}RK_{\text{global}}x \leq \frac{1}{a}x^\top Px =: \bar{a}_V\sigma(x)$, where

$$P = 10^4 \cdot \begin{bmatrix} 5.31 & 0.35 & -2.62 & 0.96 \\ 0.35 & 0.03 & -0.20 & 0.05 \\ -2.62 & -0.20 & 2.97 & -0.24 \\ 0.96 & 0.05 & -0.24 & 0.27 \end{bmatrix}. \quad (3.53)$$

We calculate $a_W = \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} = 1.40 \cdot 10^{-5}$ and $\bar{a}_V = \frac{1}{a'} = 0.0504$ and SA2 holds with $W = \bar{a}_W = 0$, similarly as in Proposition 3.7. Take $c_{\text{stop}}(\varepsilon, x) = \varepsilon x^\top Px$ with $\varepsilon > 0$ for any $x \in \mathbb{R}^n$, hence SA3 holds. Moreover Assumption 3.1 holds with $\theta(|\varepsilon|, \sigma(x)) = c_{\text{stop}}(\varepsilon, x)$. Therefore Corollary 3.2 and Theorem 3.6 follow by taking $\varepsilon \in (0, \frac{a_W}{\bar{a}_V})$, where $\frac{a_W}{\bar{a}_V} = 2.77 \cdot 10^{-4}$. Hence, for inputs of system (3.50) given by

$$\kappa = g(u, x) = \begin{cases} -[0.6, 0.6, -0.7, -0.2]x & \text{when } u = 1 \\ -[3.6, 0.9, -1.5, 0.3]x & \text{when } u = 2 \end{cases} \quad (3.54)$$

we can utilize $\text{OP}_{\text{min,stop}}$ to unite both controllers which, given an appropriate choice of ε , will calculate $u_k \in \{1, 2\}$ that preserves the global stability from input $g(2, x_k)$ while having the option to utilize the locally optimal controller $g(1, x_k)$.

We choose $\varepsilon = 10^{-9}$. In Figure 3.2, we simulate the closed-loop system with initial state $x_0 = (10, -10, -10, -10)$, with input κ either given by $\text{OP}_{\text{min,stop}}$, the local controller or

^{†7} By adding $a(Q + K_{\text{global}}^\top RK_{\text{global}})$, where $a > 0$ is a decision variable, to the block $-P$ of the LMI from [20, Theorem 2]

the global controller. We observe that: 1) the local controller indeed only locally stabilizes the origin, as it fails to converge; 2) $OP_{\min, \text{stop}}$ prioritizes the local controller, but opts for the global stabilizing one when necessary, see Figure 3.3. In Table 3.3, we compare the running cost of $OP_{\min, \text{stop}}$ versus the local and global controllers, approximately calculated with an horizon $H = 300$. We observe that $OP_{\min, \text{stop}}$ outperforms both the local and global controller.

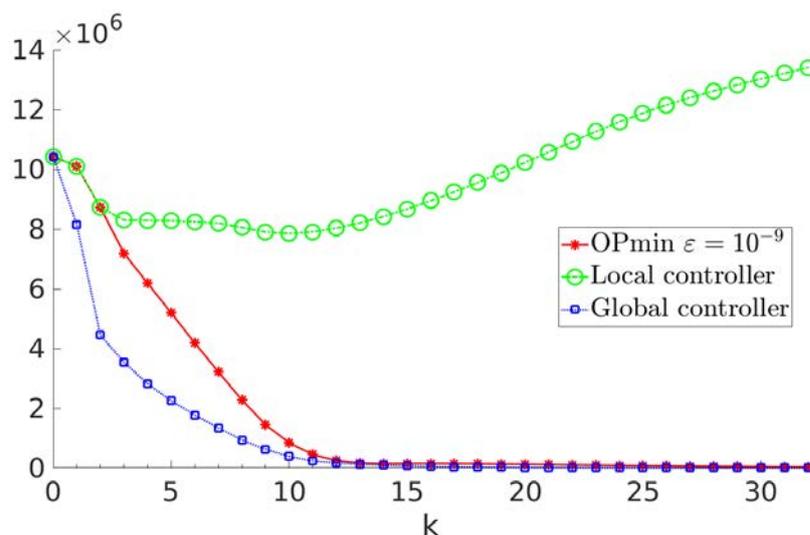


Figure 3.2: $\sigma(\phi(\cdot, x))$ for inputs given by $OP_{\min, \text{stop}}$, $g(1, x)$ and $g(2, x)$.

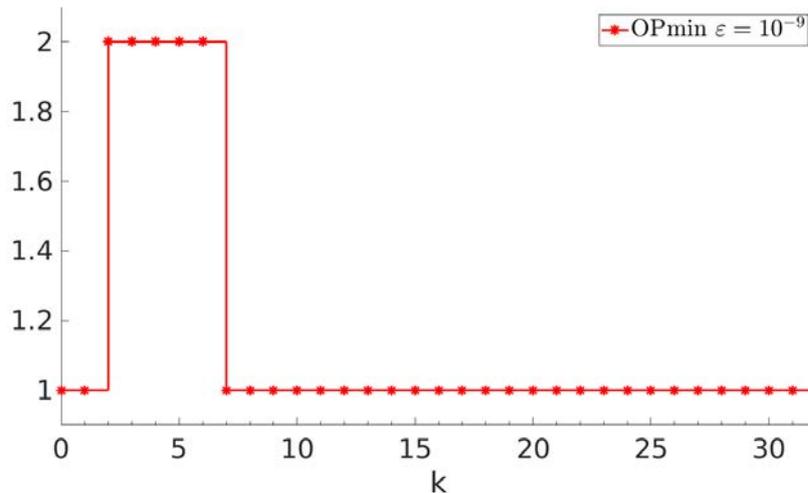


Figure 3.3: Input selection of $OP_{\min, \text{stop}}$ for $x_0 = (10, -10, -10, -10)$.

		Controller		
		OP _{min,stop}	Local	Global
	$10 \cdot (1, -1, -1, -1)$	70453	∞	83365
Initial	$10 \cdot (1, 1, 1, 1)$	15714	∞	16046
States	$(1, 1, 1, 1)$	49.75	49.75	203.77
	$(-1, -2, -3, -4)$	682.48	682.48	791.3

Table 3.3: Estimated running cost for the different controllers and various initial conditions. The symbol ' ∞ ' implies that the state is not converging towards the origin, hence the cost explodes. The minimum of each line is given in bold.

3.7 Conclusion

We have proposed a planning-based approach for the near-optimal, stable control of general nonlinear switched discrete-time systems where the control input is the switching signal. We have revisited and tailored for this purpose the optimistic planning algorithm in [32], in two different flavors. We relied on general stabilizability and detectability assumptions, originally stated in the MPC literature [24]. We have then provided near-optimality guarantees for these algorithms. Our near-optimality guarantees have major features over the bound in [32], for example it accepts the undiscounted case. We have also shown that a system controlled in a receding-horizon fashion by OP_{min} satisfies stability properties. Moreover, we have analyzed the mismatch between the optimal value function and the obtained running cost, and the same benefit as for the near-optimality guarantees were observed. The results have been applied to solve a uniting controller problem, in which we want to optimally select a controller among a bank to stabilize the origin of the plant. We have illustrated our approach using two examples.

We think that this chapter conveys an important message. It illustrates how an optimal algorithm from a different research field, namely artificial intelligence, can be adapted and tailored to solve a difficult control problem, here the near-optimal control of nonlinear

switched discrete-time systems. By building on top of the results from Chapter 2, it also demonstrates how control requirements, like stabilizability, detectability and stability, can be exploited to improve the original near-optimality guarantees of the algorithm, as we hoped for in Chapter 1.8.

From here on, there are multiple open issues to be tackled. However, before moving on, we will take a brief detour in the next chapter, in which we will re-interpret the Lyapunov property for optimal control that transpires over this manuscript. We think there is an important interpretation to be given, which is not obvious at first.

Chapter 4. A regret perspective

4.1	Introduction	95
4.2	The notion of regret	96
4.3	Examples of near-optimal control	103
4.4	Conclusion	105

4.1 Introduction

We propose in this chapter to revisit the analysis carried out in Chapters 2 and 3 in a unified way^{†1}, by exploiting the notion of what we call “regret”. Regret often appears in ADP literature [32] and RL literature [76] where it refers to the difference between the optimal value function and the obtained approximation, like e.g. Theorem 2.3 in Chapter 2.4 for instance. Here, the notion of regret is different: it is related to the (cost) consequences of “regrettable” actions. This new notion appears to be intuitive and to help a lot understanding the mechanisms, which allow concluding about stability and near-optimality.

This chapter is exploratory, and is presented as a complement of the previous chapters. No publications are associated to it.

^{†1} However, we will not enter into the same level of technical details.

4.2 The notion of regret

4.2.1 Definition

Consider again the system

$$x^+ = f(x, u), \quad (4.1)$$

with state $x \in \mathbb{R}^n$, input $u \in \mathcal{U}(x) \subseteq \mathbb{R}^m$, where $\mathcal{U}(x)$ is the nonempty set of admissible inputs for state x , and $f : \mathcal{W} \rightarrow \mathbb{R}^n$ where $\mathcal{W} := \{(x, u) : x \in \mathbb{R}^n, u \in \mathcal{U}(x)\}$. We use $\phi(k, x, \mathbf{u}|_k)$ to denote the solution to system (4.1) at time $k \in \mathbb{Z}_{\geq 0}$ with initial condition x and inputs $\mathbf{u}|_k = [u_0, u_1, \dots, u_{k-1}]$, with the convention $\phi(0, x, \cdot) = \phi(0, x, \emptyset) = x$.

We study *undiscounted* infinite-horizon costs of the form

$$J_\infty(x, \mathbf{u}) := \sum_{k=0}^{\infty} \ell(\phi(k, x, \mathbf{u}|_k), u_k) \quad (4.2)$$

where $x \in \mathbb{R}^n$, \mathbf{u} is an infinite sequence of admissible inputs and $\ell : \mathcal{W} \rightarrow \mathbb{R}_{\geq 0}$.

As in previous chapters, we assume that for any $x \in \mathbb{R}^n$, there is a sequence \mathbf{u} that minimizes cost (4.2), as formalized next.

Standing Assumption (SA). For any $x \in \mathbb{R}^n$, there exists at least one sequence of admissible inputs $\mathbf{u}_\infty^*(x)$, which is optimal, in the sense that

$$J_\infty(x, \mathbf{u}_{\gamma, d}^*(x)) = \min_{\mathbf{u}} J_\infty(x, \mathbf{u}) =: V_\infty(x), \quad (4.3)$$

where V_∞ is the *optimal cost function*. □

Because we do not know \mathbf{u}_∞^* in general, it is natural to ask how much loss of performance we might endure by taking actions that are different from optimal ones. We introduce for this purpose the regret function, which is defined as follows.

Definition 4.1. We call regret for system (4.1) and cost function (4.3), the function $R : \mathcal{W} \rightarrow \mathbb{R}_{\geq 0}$ defined as, for any $(x, u) \in \mathcal{W}$.

$$R(x, u) := \ell(x, u) + V_\infty(f(x, u)) - V_\infty(x). \quad (4.4)$$

□

The regret of action u at x is the total exceeding cost that ensues from this suboptimal (or not) input. Note that the regret differs from the Temporal Difference (TD) Error notion [42], as we define the regret relative to infinite-horizon cost function (4.3) while [42] TD error is defined relative to a given policy cost. It follows that the regret satisfies the next properties.

Lemma 4.1. *The following hold.*

(i) For any $x \in \mathbb{R}^n$ and $u \in \mathcal{U}(x)$,

$$R(x, u) \geq 0. \quad (4.5)$$

(ii) Define, for any $x \in \mathbb{R}^n$, the set

$$\mathcal{U}_\infty(x) := \left\{ u_0 : \ell(x, u_0) + V_\infty(f(x, u_0)) = V_\infty(x) \right\}. \quad (4.6)$$

Set $\mathcal{U}_\infty(x)$ is non-empty and ensures that $R(x, u) = 0$ for $u \in \mathcal{U}_\infty(x)$, and $R(x, u) > 0$ for $u \notin \mathcal{U}_\infty(x)$. \square

Proof. Let $x \in \mathbb{R}^n$ and $u \in \mathcal{U}(x)$, $R(x, u) \geq 0$ holds since $\ell(x, u) + V_\infty(f(x, u)) - V_\infty(x) \geq 0$ by (4.3) and Bellman equation, thus item (i) of Lemma 4.1 holds. For item (ii), consider $u_0^* \in \mathcal{U}_\infty(x)$, it follows that $R(x, u_0^*) = \ell(x, u_0^*) + V_\infty(f(x, u_0^*)) - V_\infty(x) = 0$ by definition of set $\mathcal{U}_\infty(x)$. Conversely, $R(x, u) = \ell(x, u) + V_\infty(f(x, u)) - V_\infty(x) \neq 0$ for $u \notin \mathcal{U}_\infty(x)$, hence $R(x, u) > 0$ since $R(x, u) \geq 0$ holds per item (i) of Lemma 4.1. \blacksquare

Item (i) of Lemma 4.1 implies that the regret is always nonnegative. On the other hand, item (ii) of Lemma 4.1 expresses, quite naturally, that we do not “regret” an optimal action $u_0^* \in \mathcal{U}_\infty(x)$. To the contrary, for any input non-optimal input $u \notin \mathcal{U}_\infty(x)$, we “regret” something and $R(x, u) > 0$. Regret occurs as a balance of two factors. The first is the amount of *immediate* extra work done by u , akin to $\ell(x, u) > \ell(x, u_0^*)$. The second is the amount of extra work *in the future* that has to be payed back, akin to $V_\infty(f(x, u)) > V_\infty(f(x, u_0^*))$. For an extreme illustration, (i) assume that there is $u \notin \mathcal{U}_\infty(x)$ such that $V_\infty(f(x, u)) = 0$, it follows from Definition 4.1 that $R(x, u) = \ell(x, u) - V_\infty(x) \geq 0$, hence $\ell(x, u) \geq V_\infty(x)$, we did more immediate work than necessary; (ii) assume that there

is $u \notin \mathcal{U}_\infty(x)$ such that $\ell(x, u) = 0$, it follows that $R(x, u) = V_\infty(f(x, u)) - V_\infty(x) \geq 0$, hence $V_\infty(f(x, u)) \geq V_\infty(x)$, our lack of work has increased the future (and total) work to be realized. In other words, regrettable inputs may have an “immediate” extra cost, paid by $\ell(x, u)$, or a “future” (to-do) cost, due to state $f(x, u)$ having an unfavorable cost $V_\infty(f(x, u))$.

We have also, by a simple rearranging of terms, in view of (4.4), for any $(x, u) \in \mathcal{W}$,

$$V_\infty(f(x, u)) - V_\infty(x) = -\ell(x, u) + R(x, u). \quad (4.7)$$

While, at a first look, we are nonthwiser compared to (4.4), consider that we can take u such that $R(x, u) < \ell(x, u)$ for $\ell(x, u) > 0$ and 0 otherwise. This immediately implies that $V_\infty(f(x, u)) < V_\infty(x)$: the next state $f(x, u)$ has a smaller optimal cost than x . When we take actions that realizes more work than regret, $\ell(x, u) > R(x, u)$, the total remaining work is reduced, $V_\infty(f(x, u)) < V_\infty(x)$. If this goes on forever, we eventually reach states where little work is left to do, i.e. the state is close to set $\{z \in \mathbb{R}^n : V_\infty(z) = 0\}$. In control terms, (4.7) implies that we have a *weak* Lyapunov function. Moreover, if we continue to generate inputs for which the realized work is entirely regrettable, i.e. $R(x, u) = \ell(x, u)$ for any $x \in \mathbb{R}^n$, the remaining work never decreases, $V_\infty(f(x, u)) = V_\infty(x)$, we reach a trajectory where the “to-do” cost is stationary.

By having a sufficiently large decay $-\ell(x, u) + R(x, u)$, we can formalize a stronger result. First, we define system

$$x^+ = f(x, K(x)) \quad (4.8)$$

where $K(x)$ is some deterministic and admissible feedback law from x , i.e. $K(x) \in \mathcal{U}(x)$ for all $x \in \mathbb{R}^n$.

Remark 4.1. Note that we use an explicit feedback law $K(x)$ to make notation concise and clear. The following results can be relaxed for any infinite input sequence that verifies similar conditions for the pair of state at time-step k (and its respective input), for all $k \in \mathbb{Z}_{\geq 0}$. □

We denote $\phi^K(k, x)$ denote the solution for system (4.8) at time $k \in \mathbb{Z}_{\geq 0}$. If we assume, that there exist $a \in [0, 1)$ such that, for any $x \in \mathbb{R}^n$, $R(x, u) \leq a\ell(x, u)$ we show that set $\{z \in \mathbb{R}^n, V_\infty(z) = 0\}$ is reached asymptotically. We formalize this in the next result.

Proposition 4.1. *For all $x \in \mathbb{R}^n$, let feedback $K(x)$ be such that $R(x, K(x)) \leq a\ell(x, K(x))$ for some $a \in [0, 1)$. Then, for all $x \in \mathbb{R}^n$ the sequence $\{V_\infty(\phi^K(k, x))\}$ converges to 0, and the sequence is bounded by $V_\infty(x)$, i.e. $V_\infty(\phi^K(k, x)) \leq V_\infty(x)$.*

Proof. Let $x \in \mathbb{R}^n$ and $k \in \mathbb{Z}_{\geq 0}$. First, we show that the sequence of cost over time is bounded. $V_\infty(\phi^K(k+1, x)) - V_\infty(\phi^K(k, x)) \leq -(1-a)\ell(\phi^K(k, x), K(\phi^K(k, x)))$ where $\phi^K(k, x)$ is the solution of autonomous system $x^+ = f(x, K(x))$ at time $k \in \mathbb{Z}_{\geq 0}$. Since $\ell(\phi^K(k, x), K(\phi^K(k, x))) \geq 0$, $V_\infty(\phi^K(k, x)) \leq V_\infty(\phi^K(0, x))$, hence the sequence is bounded by $V_\infty(x)$. Now, we show that $V_\infty(\phi^K(k, x)) \rightarrow 0$. Clearly, for any $x \in \{x \in \mathbb{R}^n : V_\infty(x) = 0\}$, follows $V_\infty(\phi^K(k, x)) = 0$ for all $k \in \mathbb{Z}_{\geq 0}$. For $\{x \in \mathbb{R}^n : V_\infty(x) > 0\}$ we proceed by contradiction. Assume that there exist $x \in \mathbb{R}^n$ with $V_\infty(x) = \epsilon > 0$ such that $V_\infty(\phi^K(k, x)) > a\epsilon$ for all $k \in \mathbb{Z}_{\geq 0}$, that is $\sum_{n=0}^k (1-a)\ell(\phi^K(n, x)) \leq V_\infty(x) - V_\infty(\phi^K(k, x)) < (1-a)\epsilon$, hence $\sum_{n=0}^k \ell(\phi^K(n, x)) < \epsilon$. However, when $k \rightarrow \infty$, we know from SA $\lim_{k \rightarrow \infty} \sum_{n=0}^k \ell(\phi^K(n, x)) \geq V_\infty(x)$ as $\sum_{n=0}^k \ell(\phi^K(n, x))$ is (in general) a non-optimal infinite-horizon cost. Therefore $\epsilon = V_\infty(x) \leq \sum_{n=0}^k \ell(\phi^K(n, x)) < \epsilon$, a contradiction is attained. Then, for every $x \in \mathbb{R}^n$ there exist a finite k such that $V_\infty(\phi^K(k, x)) \leq aV_\infty(x)$. By iteration, we conclude that $V_\infty(\phi^K(k, x)) \rightarrow 0$ as $k \rightarrow \infty$. \blacksquare

4.2.2 Near-optimality guarantees

The notion of regret allows to easily derive properties on the running and average costs as explained below.

Running cost. Given (4.7) for any admissible feedback law $K(\cdot)$ and its associated autonomous system (4.8), we also have by simple rearranging of terms that, for any $x \in \mathbb{R}^n$,

$$\ell(x, K(x)) = V_\infty(x) - V_\infty(f(x, K(x)) + R(x, K(x))). \quad (4.9)$$

For the sake of convenience, we denote $\ell_k^\phi(x) = \ell(\phi^K(k, x), K(\phi^K(k, x)))$ the stage-cost incurred at time k and $R_k^\phi := R(\phi^K(k, x), K(\phi^K(k, x)))$ the corresponding regret. Hence

$$\ell_0^\phi(x) = V_\infty(x) - V_\infty(\phi^K(1, x)) + R_0^\phi(x). \quad (4.10)$$

We then derive for any $N \in \mathbb{Z}_{>0}$,

$$\sum_{k=0}^N \ell_k^\phi(x) = V_\infty(x) - V_\infty(\phi^K(N+1, x)) + \sum_{k=0}^N R_k^\phi(x). \quad (4.11)$$

When $\sum_{k=0}^\infty R_k^\phi(x)$ is summable, we derive the following result.

Lemma 4.2. *Assume that $\sum_{k=0}^\infty R_k^\phi(x)$ is finite for any $x \in \mathbb{R}^n$. Then $\sum_{k=0}^\infty \ell_k^\phi(x)$ is summable and the following holds*

$$V_\infty(x) \leq \sum_{k=0}^\infty \ell_k^\phi(x) \leq V_\infty(x) + \sum_{k=0}^\infty R_k^\phi(x). \quad (4.12)$$

Proof. Let $x \in \mathbb{R}^n$. Since the image of ℓ is nonnegative, $\sum_{k=0}^\infty \ell_k^\phi(x)$ is summable if the upperbound in (4.11) is finite. This is the case since $\sum_{k=0}^\infty R_k^\phi(x)$ is assumed summable^{†2} and $V_\infty(x) - V_\infty(\phi^K(N+1, x)) \leq V_\infty(x)\infty$ per virtue of SA and $-V_\infty(\phi^K(N+1, x)) \leq 0$. Hence, $\ell_k^\phi(x) \leq V_\infty(x) + \sum_{k=0}^\infty R_k^\phi(x)$. On the other hand, we derive $V_\infty(x) \leq \sum_{k=0}^\infty \ell_k^\phi(x)$ follows from SA. ■

We have derived a bound on the running cost, where the only requirements are $\ell(x, u) \geq 0$, SA, and a summable regret. Inequality (4.12) implies that the distance of the running cost to the true infinite-horizon cost is proportional to the size of the total regret as formalized next.

Proposition 4.2. *Assume $\sum_{k=0}^\infty R_k^\phi(x)$ finite for any $x \in \mathbb{R}^n$, then*

$$\left| \sum_{k=0}^\infty \ell_k^\phi(x) - V_\infty(x) \right| \leq \sum_{k=0}^\infty R_k^\phi(x). \quad (4.13)$$

□

Proposition 4.2 is meaningful. It implies that, when controlling a system with inputs issued from a near-optimal algorithm, the regret has to decrease sufficiently fast in order to guarantee that the running cost performance is close to the true optimal cost.

Average cost. When $\sum_{k=0}^\infty R_k^\phi(x)$ is not summable, we can look at the average cost along solutions, i.e. $\frac{1}{N} \sum_{k=0}^N \ell_k^\phi(x)$. We derive the following result.

^{†2} For and only if statement, we have to look at the convergence of $-V_\infty(\phi^K(N+1, x)) + \sum_{k=0}^N R_k^\phi(x)$

Lemma 4.3. Assume $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^N R_k^\phi(x)$ is finite for all $x \in \mathbb{R}^n$ and

$$\lim_{N \rightarrow \infty} \frac{1}{N} V_\infty(\phi^K(N+1, x)) = 0. \quad (4.14)$$

Then, for any $x \in \mathbb{R}^n$,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^N \ell_k^\phi(x) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^N R_k^\phi(x). \quad (4.15)$$

In addition, when $\lim_{N \rightarrow \infty} \frac{1}{N} V_\infty(\phi^K(N+1, x)) \neq 0$, then, for any $x \in \mathbb{R}^n$,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^N \ell_k^\phi(x) \leq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^N R_k^\phi(x). \quad (4.16)$$

□

Proof. Let $x \in \mathbb{R}^n$. Note that $\lim_{N \rightarrow \infty} \frac{1}{N} \left(V_\infty(x) - V_\infty(\phi^K(N+1, x)) + \sum_{k=0}^{\infty} R_k^\phi(x) \right) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^N R_k^\phi(x)$ since $\frac{1}{N} V_\infty(x) \rightarrow 0$ and $-\frac{1}{N} V_\infty(\phi^K(N+1, x)) \rightarrow 0$ as $N \rightarrow \infty$ in view of (4.14). Hence (4.15) follows from (4.11) divided by N . Furthermore, inequality (4.16) follows similarly from (4.11) and $-V_\infty(\phi^K(N+1, x)) \leq 0$. ■

In sum, the average cost is simply the average regret. When the regret converges, but not to 0, the system will potentially observe stage-costs (at average) as big as the regret (at average).

Comparison to ADP. We can relate the above notion to the error-bounds usually derived in ADP literature for particular algorithms. We stress that they are not the same, as the notion of regret is explicitly related to the autonomous system (4.8), while in ADP, in general, is centered around the near approximation of cost V_∞ .

We illustrate this difference in the following manner. Assume there is an estimate \widehat{V} of V_∞ , which satisfies the next property.

Assumption 4.1. There exists $\epsilon > 0$ such that for any $x \in \mathbb{R}^n$,

$$0 \leq \widehat{V}(x) - V_\infty(x) \leq \epsilon. \quad (4.17)$$

□

Assume also that there is an associated stationary policy $\widehat{u}(x)$, with associated cost $\widehat{\ell}(x) := \ell(x, \widehat{u}(x))$, that maps to state $\widehat{v} := f(x, \widehat{u}(x))$. Thus, we derive that the regret satisfies, in view of (4.17),

$$R(x, \widehat{u}(x)) \leq \epsilon + V_\infty(\widehat{v}) + \widehat{\ell}(x) - \widehat{V}(x). \quad (4.18)$$

The regret can be related to the near-optimal error ϵ , however it also depends on $V_\infty(\widehat{v}) + \widehat{\ell}(x) - \widehat{V}(x)$, which cannot be bounded by Assumption 4.1 without additional assumptions. In sum, regret provides different information than (4.17).

4.2.3 Stability guarantees

We now revisit the stability analysis in Chapters 2 and 3 based on the notion of regret, under the same assumptions.

Assumption 4.2. There exist $\bar{\alpha}_V, \alpha_W \in \mathcal{K}_\infty$, continuous functions $W, \sigma : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$, $\bar{\alpha}_W : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ continuous, non-decreasing and zero at zero, such that the following conditions hold.

(i) For any $x \in \mathbb{R}^n$,

$$V_\infty(x) \leq \bar{\alpha}_V(\sigma(x)). \quad (4.19)$$

(ii) For any $x \in \mathbb{R}^n$,

$$W(x) \leq \bar{\alpha}_W(\sigma(x)) \quad (4.20)$$

$$W(f(x, K(x)) - W(x) \leq -\alpha_W(\sigma(x)) + \ell(x, K(x)). \quad (4.21)$$

□

Note that instead of stating item (ii) of any $u \in \mathcal{U}(x)$, we only consider input $K(x)$.

We are ready to state the following global asymptotic stability result.

Proposition 4.3. *Suppose Assumption 4.2 holds, and let, for all $x \in \mathbb{R}^n$, feedback $K(x)$ be such $R(x, K(x)) \leq a\alpha_W(\sigma(x))$ for some $a \in [0, 1)$. Then, for all $x \in \mathbb{R}^n$ any solution $\phi^K(\cdot, x)$ to system (2.5) satisfies, for all $k \in \mathbb{Z}_{\geq 0}$*

$$\sigma(\phi^K(k, x)) \leq \max\{\beta(\sigma(x), k)\}. \quad (4.22)$$

□

Sketch of proof. Let $x \in \mathbb{R}^n$. We follow the steps of the proof of Theorem 3.4. We have then that $Y_\infty := V_\infty(x) + W(x)$ verifies, for all $x \in \mathbb{R}^n$,

$$(i) \quad \alpha_W(\sigma(x)) \leq Y_\infty(x) \leq (\bar{\alpha}_V + \bar{\alpha}_W)(\sigma(x))$$

$$(ii) \quad Y_\infty(x^+) - Y_\infty(x) \leq -\alpha_W(\sigma(x)) + R(x, u), \text{ where } x^+ = f(x, K(x)).$$

Given $R(x, K(x)) \leq a\alpha_W(\sigma(x))$ for any $x \in \mathbb{R}^n$, it follows that Y_∞ is a Lyapunov function that guarantees global asymptotic stability for the set $\{x \in \mathbb{R}^n : \sigma(x) = 0\}$. ■

In Proposition 4.3 we have converted the convergence to set $\{x : V_\infty(x) = 0\}$ in Proposition 4.1 to the stability of set $\{x : \sigma(x) = 0\}$. Note that $R(x, K(x)) \leq a\alpha_W(\sigma(x))$ does not imply $R(x, K(x)) \leq a\ell(x, u)$. Indeed, given that the detectability assumption is verified with $W(x) > 0$ for all $x \in \{z \in \mathbb{R}^n : \sigma(z) > 0\}$, then $\ell(x, u)$ might be 0 while $\alpha_W(\sigma(x)) > 0$. In fact, we have $R(x, u) \leq a(W(x) - W(f(x, u)) + \ell(x, u))$. Compared to Proposition 4.1, Proposition 4.3 shows that stability has different regret requirements than just reducing the amount of remaining work.

Remark 4.2. As seen in Chapter 2.3.2 and 3.4.2, semiglobal practical stability can be ensured in general. Semiglobal practical stability results can be derived by relaxing the requirement that $R(x, K(x)) \leq a\alpha_W(\sigma(x))$ assumption in a weaker form. □

We now illustrate our notion of regret on a series of examples.

4.3 Examples of near-optimal control

4.3.1 An unstable optimal feedback law

Consider an unstable scalar plant^{†3},

$$x^+ = ax + u,$$

^{†3} This system and cost was also considered in [54].

where $x, u \in \mathbb{R}$ and $a > 1$. Let the stage-cost be $\ell(x, u) = u^2$, clearly the optimal sequence is $\mathbf{u}_\infty^* = [0, 0, \dots]$, which makes the origin trivially unstable. However, any controller $K(x) = (-a + b)x$ for $b \in (-1, 1)$ is stable,

$$x(k+1) = b^k x.$$

The regret is well-defined and given by

$$R(x, K(x)) = \ell(x, K(x)) = (-a + b)^2 x^2.$$

Every action is “regrettable” in the sense that $\ell(x, K(x))$ is always equal to $R(x, K(x))$ for any $x \in \mathbb{R}^n$. However stability is still ensured, and we can exactly calculate the running cost by invoking (4.12), that is $V^{\text{run}}(x) = \frac{(-a+b)^2}{1-b^2} x^2$.

4.3.2 Finite-horizon optimal feedback law

In this subsection, we recover the results from Chapters 2 and 3 by showing that finite-horizon optimal costs are viable strategies to upper-bound the regret.

Fixed finite-horizon. Let

$$\hat{v} = f(x, \hat{u}),$$

where \hat{u} is the first input of sequence $\mathbf{u}_d^*(x)$ such that $V_d(x) = J_d(x, \mathbf{u}_d^*(x))$, as in Chapter 2 with no discount factor, i.e. $\gamma = 1$. Given SA and item (i) of Assumption 4.2 and by following the steps of the proof of Theorem 3.4, we derive that

$$V_\infty(\hat{v}) \leq V_\infty(x) - \ell(x, \hat{u}) + \bar{\alpha}_V(\sigma(\phi_d)) \quad (4.23)$$

where $\phi_d(x) = \phi(d, x, \mathbf{u}_d^*(x)|_d)$ denotes the last state reached by the d -horizon optimal sequence. Thus,

$$R(x, \hat{u}) = \ell(x, \hat{u}) + V_\infty(\hat{v}) - V_\infty(x) \leq \bar{\alpha}_V(\sigma(\phi_d(x))). \quad (4.24)$$

The regret is bounded by the size of the last state of an finite-horizon optimal sequence.

When item (ii) of Assumption 4.2 holds for any input $u \in \mathcal{U}(x)$, like in item (ii) of Assumption 2.1, we have that $\bar{\alpha}_V(\sigma(\phi_d(x)))$ can be made as small as desired by having

$d \rightarrow \infty$, see Proposition 3.3. As the regret bound is controlled by the horizon d , we can then derive cost and stability guarantees, as underlined in Sections 4.2.2 and 4.2.3.

Minimum guaranteed horizon. Similarly, in $\text{OP}_{\text{min,budget}}$, we have

$$\hat{v} = f(x, u_B),$$

where u_B is the first input of sequence $\mathbf{u}_{d(x)}^*(x)$ such that $V_{d(x)}(x) = J_{d(x)}(x, \mathbf{u}_{d(x)}^*(x))$, and $\phi_{d(x)}(x) = \phi(d(x), x, \mathbf{u}_{d(x)}^*(x)|_{d(x)})$. In Chapter 3, we have guaranteed that $d(x) \geq \bar{d}$ for some $\bar{d} \geq 0$, and moreover that $V_\infty(\hat{v}) \leq V_\infty(x) - \ell(x, \hat{u}) + \bar{\alpha}_V(\sigma(\phi_{d(x)}(x)))$. Hence,

$$R(x, u_B) \leq \bar{\alpha}_V(\sigma(\phi_{\bar{d}})),$$

where $\sigma(\phi_{\bar{d}})$ is an upper-bound on measure $\sigma(\phi_{d(x)})$. By having a large enough computational budget B , we have $\bar{d} \rightarrow \infty$, which in turn implies $\sigma(\phi_{\bar{d}})$ small. Since we can make the regret as small as desired, we can then endow some stability and cost guarantees like in the fixed finite-horizon case.

Controlling the last state measure. For $\text{OP}_{\text{min,stop}}$, the stopping criterion guarantees $\sigma(\phi_{d(x)}(x)) \leq c_{\text{stop}}(\varepsilon, x)$, hence

$$R(x, u_{\text{stop}}) \leq \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x)).$$

Similar to previous cases, we can make the regret as small as desired, here by correctly choosing $c_{\text{stop}}(\varepsilon, x)$, and derive stability and near-optimality guarantees.

Note that in all four cases, we require no particular information on how input $K(x)$ is generated, e.g. the underlying algorithm is not exploited. Only the bound for $R(x, K(x))$ and an associated parameter to reduce its size is sufficient to endow the corresponding closed-loop system some stability and cost properties.

4.4 Conclusion

We have introduced the notion of regret, which seems to be an instructive tool to explain the inner-workings of the relationship between near-optimality, closed-loop stability, and

running cost guarantees. Using the regret, we are able to redescribe the results from Chapters 2 and 3, where the near-optimality term provides a bound for the regret. This let us quickly recoverer stability, running cost or average cost guarantees in a more unified manner. In one example, we are able to describe a stable sequence with a well defined running cost, despite the fact that the infinite-horizon optimal solution diverges.

Chapter 5. Conclusions

5.1 Overview

In this thesis, we have investigated the stability of nonlinear discrete-time systems controlled by inputs that minimize a discounted cost in general. We have exploited techniques from model predictive control [24] as well as advanced nonlinear tools for Lyapunov stability for this purpose. An important consequence of stability is that we can revisit and improve the near-optimality analysis for algorithms from the Approximate Dynamic Programming (ADP) literature. Our results apply in particular to (Approximated) Value Iteration (AVI) [7], an important pillar of ADP literature, and Optimistic Planning (OP) [32], so it is better suited and natural for control engineers. We now briefly review the contributions chapter by chapter.

In Chapter 2, we have examined stability of nonlinear systems driven by inputs that minimize a finite-horizon discounted cost. This covers VI as a particular case, as explained in Chapter 1. Under classical assumptions in a control context (namely stabilizability and detectability), we have shown that stability is guaranteed when both the discount factor *and* the horizon, which corresponds to the number of iterations for VI, are sufficiently close to 1 and sufficiently large, respectively. In general, only semiglobal practical stability is guaranteed. Semiglobal asymptotic and exponential stability results are also provided under stronger assumptions. We then show that a discounted finite-horizon optimal problem can be made as close as wanted to the infinite-horizon discounted cost, by taking the discount factor and horizon sufficiently close to 1 and ∞ , respectively. These results extend to the case where inputs are only near-optimal for the finite-horizon optimal control

problem. Hence, we have successfully emulated the approach from AVI of using a finite-horizon as a proxy of the infinite-horizon control problem, providing not only near-optimal guarantees but also stability properties. The provided near-optimal guarantees differ from those typically found in the ADP literature. We have shown that the finite-horizon optimal cost can be made as close as desired to the infinite one by making the horizon larger, while allowing the undiscounted case. Our results are illustrated on an example of an AVI scheme, which corroborated the theory.

In Chapter 3, we have concentrated on a more specific optimal control algorithm from the ADP literature, Optimistic Planning (OP) [32]. OP also uses a finite-horizon problem as a proxy to the infinite-horizon one, however it is specialized for nonlinear systems with a finite number of possible inputs. It is thus relevant for switched systems with nonlinear modes for which the switching signal is the control input; an important class of systems in the control literature. We have proposed a modified algorithm called OP_{\min} that relaxes some assumptions to be more aligned to our desire for stability. OP_{\min} was displayed in two forms: one with a fixed number of calculations, for applications where a fixed running time is important for instance; and another with a parameterized stopping criterion, which has stronger stability guarantees and less conservative near-optimality estimates for similar computational budgets, however the running time is not fixed. Like in Chapter 2, we have provided near-optimal and stability guarantees for these algorithms. The stability results of this chapter extend those of Chapter 2 to accept a variable, state dependent, horizon. In addition, we have provided running cost guarantees for such algorithms, which is the actual cost incurred by the autonomous system. Similarly to Chapter 2, the provided near-optimal guarantees differ from those typically found in the ADP literature, and we have shown that we can tune OP_{\min} to better approximate the infinite horizon optimal cost as desired, while allowing the undiscounted case. This is also true for the running cost, that is it can be made as close to the optimal infinite-horizon sequence as desired, at a cost of more computations.

In Chapter 4, we have revisited the technical aspects for near-optimal control and stability of the previous chapters. To accept the variable optimization horizon from OP_{\min}

in Chapter 3, we have utilized the infinite-horizon optimal cost as basis for our Lyapunov function. In contrast, Chapter 2 and previous literature [24] construct a different Lyapunov function for each horizon, with the finite-horizon optimal cost as basis. By fixing the Lyapunov function to be the infinite-horizon cost (with the addition of a suitable term related to the detectability property of the system), we have derived the notion of “regret”, which has let us re-derive the technical aspects of this manuscript in a concise form.

The results of this thesis allow to envision a range of promising future research directions, as listed in the sequel.

5.2 Perspectives

In the following, we list extensions of the presented work that are viable under a proof-of-concept and some other ideas we believe to be scientifically relevant but do require more work. We classify the former as “short term”, and the latter as “longer term”.

5.2.1 Short term

A more general detectability condition. We have considered in this work a more restrictive form of detectability compared to the one in [24] and [55]. The discount factor is a major source of technicalities for stability under this more general form of detectability. Still, we are convinced the results of this thesis can be extended to this case, for which a compelling example would be welcome.

Other ADP algorithms. ADP is vast in algorithms, we have only considered a particular form of AVI in Chapter 2 and the more niche algorithm OP in Chapter 3. We are convinced that the general approach presented in this manuscript can be used to address other ADP algorithms. In particular, policy iteration, a main pillar of ADP literature, was left for future work, and is highly relevant.

Tailored proofs. While the results in Chapter 2 are applicable to problems at large, we expect that tailored approaches, like the one for OP_{\min} in Chapter 3, to be of major importance. Not only we might derive tailored stability properties and derive less conservative

estimates of the discount factor or of the horizon, but we hope to modify algorithms to provide stability upfront. For example, an ongoing work [56] is presented in Appendix D, provides tailored assumptions and stability analysis for AVI. Research in this direction is also a good exercise in learning the strengths and weaknesses of each particular ADP tool, and an opportunity to relate these in terms of stability.

“Under budget” OP_{\min} . OP_{\min} in Chapter 3 implicitly requires that we have a sufficient computational budget at any time instant. This may not be the case in embedded systems, for which the resources are limited and may be shared with other tasks. We have thus started to investigate the scenario where we may not always be able to minimize a long-enough horizon cost at each step, which we model as an optimization problem with a time-varying horizon, where the latter depends on the time and not the state as in Chapter 3. The idea is that we should be able to tolerate very short horizons sporadically (i.e. few computations) as long as these are compensated by longer horizons afterwards. We seek to analyse the resulting impact on stability and optimality guarantees, and if necessary, how to adapt OP_{\min} on-the-fly to recover (or be robust) to such events.

Optimizing uniting control. In Chapter 3.6 we initiated a study of the optimizing uniting control problem. We are keen in having a more in-depth analysis, with relaxed assumptions, like removing the assumption about the existence of a global stabilizing controller. A closely related topic, but not limited to, is whether OP_{\min} scales well for $M \gg 2$ controllers.

5.2.2 Longer term

MPC contributions 1. It is common in MPC to adopt terminal costs for the optimal control problem. For example, this is done in [24] to drastically reduce the horizon. Since this changes the cost to be optimized, it might impact the near-optimal analysis, in particular the running cost. On the other hand, VI naturally implements terminal costs when initiated with a cost function different than 0. Hence, a logical extension of our results is to accept terminal costs, which we expect will better align MPC and ADP tools objectives.

MPC contributions 2. The same could be said to results that consider state constraints, which are also commonplace in MPC literature and are of fundamental importance for safety reasons.

Stronger robustness guarantees. Another main subject of interest is dealing with errors in the optimal control problem due to model uncertainties. We have provided some results in this direction, like robustness properties for the closed-loop system, but only in a qualitative sense. Given how essential robustness is, we believe a more complete story in this direction to be relevant.

Link between robustness and model approximations. Approximations in the optimal control problem happen either due to imperfect algorithms or model uncertainties. In Chapter 2, we have shown that stability and optimality guarantees holds for near-optimal finite-horizon optimal control, as an extension to the exact case. However, we have not dwelt deeper at the link between robustness and approximation errors. Notably, the near-optimal assumption in Chapter 2, which accepts imperfect algorithms, also encompasses certain model uncertainties. This seems to corroborate the intuitive reasoning that solving for larger horizons is problematic when the model diverges from the true system, as we expect to observe larger accumulated errors. Whether a more explicit link could be described is a relevant question.

OP_{min} for linear switched systems. Several other research tracks are available for OP_{min}. We would like to study the potential of OP_{min} for the particular case of linear switched systems with quadratic stage costs and to compare its performance with respect to the related literature, e.g. [3, 62, 78].

Stochastic dynamics. Another path of interest would be to provide a stochastic version of this thesis. To guide such extension, we could provide a stochastic version of OP_{min} inspired by [14].

OP_{min} for quantized control. Other less concrete but interesting points for OP_{min} are, for example: to modify more extensively the algorithm to reduce branching, to relax

the requirement of solving a finite-horizon problem exactly, to study the viability of OP_{\min} for a quantized scenario, i.e. when we have flexibility in defining the inputs in a some compact set.

Exploiting the notion of regret. Finally, there is the question of whether this emergent concept of regret in Chapter 4 is a useful tool to derive stability and running cost guarantees for other near-optimal derivations for the infinite-horizon control problem.

Stability guarantees for Reinforcement Learning. Given that DP is at the kernel of RL, and since we now have strong tools to analyse the stability of systems controlled by DP methods, it will be interesting to see how they extend to RL case.

Appendix A. Technical results for Chapter 2

The next Lyapunov properties are used in the proof of Theorem 2.1.

Theorem A.1. *Suppose Assumption 2.1 is verified. Functions $\underline{\alpha}_Y, \bar{\alpha}_Y, \alpha_Y \in \mathcal{K}_\infty$ from Theorem 2.1 are such that, for any $\gamma \in (0, 1]$, $d \in \mathbb{Z}_{>0}$, $k \in \{0, \dots, d\}$, and $Y_{\gamma, d-k} := V_{\gamma, d-k} + W$ is such that the following holds.*

(i) For any $x \in \mathbb{R}^n$,

$$\underline{\alpha}_Y(\sigma(x)) \leq Y_{\gamma, d-k}(x) \leq \bar{\alpha}_Y(\sigma(x)). \quad (\text{A.1})$$

(ii) For any $x \in \mathbb{R}^n$,

$$\begin{aligned} & Y_{\gamma, d-(k+1)}(\phi_{k+1}^*) - Y_{\gamma, d-k}(\phi_k^*) \\ & \leq \frac{1}{\gamma} \left(-\alpha_Y(\sigma(\phi_k^*)) + (1 - \gamma)Y_{\gamma, d-k}(\phi_k^*) \right), \end{aligned} \quad (\text{A.2})$$

where $\phi_{k+1}^* \in F_{\gamma, d-k}^*(\phi_k^*)$ and $\phi_0^* = x$. □

Proof of Theorem A.1.

Let $\gamma \in (0, 1]$, $d \in \mathbb{Z}_{>0}$, $k \in \{0, \dots, d\}$, $x \in \mathbb{R}^n$. There exists $[u_0^*, u_1^*, \dots, u_d^*] = \mathbf{u}_{\gamma, d}^*(x)$ where $\mathbf{u}_{\gamma, d}^*(x)$ is an optimal input sequence for system (2.1) with cost (2.2) according to SA.

Define $\ell_0^* = \ell(\phi_0^*, u_0^*)$ and let \mathbf{u}' be the sequence of length d such that $\mathbf{u}_{\gamma, d}^*(x) = [u_0^*, \mathbf{u}']$. From the definition of $V_{\gamma, d}$ in (2.3) and cost (2.2), $V_{\gamma, d}(x) = J_{\gamma, 0}(x, u_0^*) + \gamma J_{\gamma, d-1}(f(x, u_0^*), \mathbf{u}') = \ell_0^* + \gamma J_{\gamma, d-1}(\phi_1^*, \mathbf{u}')$ By definition of $V_{\gamma, d}(x)$ and \mathbf{u}' , $V_{\gamma, d-1}(\phi_1^*) =$

$J_{\gamma,d-1}(f(x, u_0^*), \mathbf{u}')$. Hence, $V_{\gamma,d}(x) = \ell_0^* + \gamma V_{\gamma,d-1}(\phi_1^*)$. Let $\ell_k^* = \ell(\phi_k^*, u_k^*)$. Remark that, when $k < d$ and by iteration,

$$\begin{aligned} V_{\gamma,d-k}(\phi_k^*) &= J_{\gamma,0}(\phi_k^*, u_k^*) + \gamma J_{\gamma,d-(k+1)}(f(\phi_k^*, u_k^*), \mathbf{u}_{\gamma,d-(k+1)}^*(f(\phi_k^*, u_k^*))) \\ &= \ell_k^* + \gamma \min_{\mathbf{u}} J_{\gamma,d-(k+1)}(\phi_{k+1}^*, \mathbf{u}) \\ &= \ell_k^* + \gamma V_{\gamma,d-(k+1)}(\phi_{k+1}^*). \end{aligned} \quad (\text{A.3})$$

Notice that for $k = d$, $V_{\gamma,0}(\phi_d^*) = \ell_d^*$ since $J_{\gamma,0}(x, u) = \ell(x, u)$. Since stage cost ℓ is nonnegative and item (i) of Assumption 2.1 holds for any d , it follows from (A.3)

$$\ell_k^* \leq V_{\gamma,d-k}(\phi_k^*) \leq \bar{\alpha}_V(\sigma(\phi_k^*)). \quad (\text{A.4})$$

Furthermore, using Bellman equation,

$$\begin{aligned} V_{\gamma,d-(k+1)}(\phi_{k+1}^*) - V_{\gamma,d-k}(\phi_k^*) &\leq V_{\gamma,d-(k+1)}(\phi_{k+1}^*) - \ell_k^* - \gamma V_{\gamma,d-(k+1)}(\phi_{k+1}^*) \\ &\leq -\ell_k^* + (1 - \gamma)V_{\gamma,d-(k+1)}(\phi_{k+1}^*). \end{aligned} \quad (\text{A.5})$$

It follows from (A.3) that $V_{\gamma,d-(k+1)}(\phi_{k+1}^*) = \frac{1}{\gamma}(V_{\gamma,d-k}(\phi_k^*) - \ell_k^*)$, hence

$$V_{\gamma,d-(k+1)}(\phi_{k+1}^*) - V_{\gamma,d-k}(\phi_k^*) \leq \frac{1}{\gamma}(-\ell_k^* + (1 - \gamma)V_{\gamma,d-k}(\phi_k^*)). \quad (\text{A.6})$$

On the other hand, in view of item (ii) of Assumption 2.1, $W(f(\phi_k^*, u_k^*)) - W(\phi_k^*) \leq -\alpha_W(\sigma(\phi_k^*)) + \ell(\phi_k^*, u_k^*)$. Since $\phi_{k+1}^* = f(\phi_k^*, u_k^*)$ and $\ell_k^* = \ell(\phi_k^*, u_k^*)$, $W(\phi_{k+1}^*) - W(\phi_k^*) \leq -\alpha_W(\sigma(\phi_k^*)) + \ell_k^*$. Furthermore, since $\gamma \leq 1$, $\gamma W(\phi_{k+1}^*) - W(\phi_k^*) \leq -\alpha_W(\sigma(\phi_k^*)) + \ell_k^*$, hence $W(\phi_{k+1}^*) - W(\phi_k^*) - \frac{1-\gamma}{\gamma}W(\phi_k^*) \leq -\frac{\alpha_W(\sigma(\phi_k^*))}{\gamma} + \frac{\ell_k^*}{\gamma}$ as in (2.19). Therefore,

$$W(\phi_{k+1}^*) - W(\phi_k^*) \leq \frac{1}{\gamma}(-\alpha_W(\sigma(\phi_k^*)) + \ell_k^* + (1 - \gamma)W(\phi_k^*)) \quad (\text{A.7})$$

We define $Y_{\gamma,k} := V_{\gamma,k} + W$. In view of (A.6) and (A.7),

$$\begin{aligned} Y_{\gamma,d-(k+1)}(\phi_{k+1}^*) - Y_{\gamma,d-k}(\phi_k^*) &\leq \frac{1}{\gamma}(-\ell_k^* + (1 - \gamma)V_{\gamma,d-k}(\phi_k^*) - \alpha_W(\sigma(\phi_k^*)) + \ell_k^* \\ &\quad + (1 - \gamma)W(\phi_k^*)) \\ &= \frac{1}{\gamma}(-\alpha_W(\sigma(\phi_k^*)) + (1 - \gamma)Y_{\gamma,d-k}(\phi_k^*)). \end{aligned} \quad (\text{A.8})$$

Thus item (ii) is verified with $\alpha_Y = \alpha_W \in \mathcal{K}_\infty$.

On the other hand, item (i) of Theorem A.1 follows by noting that item (i) of Theorem 2.1 holds for any $x \in \mathbb{R}^n$, $d \in \mathbb{Z}_{>0}$ and $\gamma \in (0, 1]$. ■

Appendix B. Technical results for Chapter 3

B.1 Auxiliary Result for Theorem 3.4	115
B.2 Semiglobal asymptotic stability property for $\text{OP}_{\min, \text{stop}}$. . .	116
B.3 Average cost and cost-to-attractor guarantees for $\text{OP}_{\min, \text{stop}}$	118

B.1 Auxiliary Result for Theorem 3.4

We show that, for any $d \in \mathbb{Z}_{>0}$ and $x \in \mathbb{R}^n$, the following properties hold for any finite-horizon optimal sequence $\mathbf{u}_d^*(x)$.

Lemma B.1. *Let $x \in \mathbb{R}^n$. For any $d \in \mathbb{Z}_{>0}$ and $\mathbf{u}_d^*(x)$, the following hold.*

(i) $\sigma(\phi(d, x, \mathbf{u}_d^*(x)|_d)) \leq \alpha_W^{-1} \circ (\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1})^{(d)} \circ \bar{\alpha}_Y(\sigma(x))$, with $\bar{\alpha}_Y = \bar{\alpha}_V + \bar{\alpha}_W$ and $\alpha_W, \bar{\alpha}_V, \bar{\alpha}_W \in \mathcal{K}_\infty$ comes from SA2.

(ii) For any $s > 0$, $(\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1})^{(d)}(s) < (\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1})^{(d-1)}(s)$.

(iii) For any $s \geq 0$, $\lim_{d \rightarrow \infty} (\mathbb{I} - \alpha_W \circ \bar{\alpha}_Y^{-1})^{(d)}(s) = 0$.

Proof. Let $x \in \mathbb{R}^n$ and $d \in \mathbb{Z}_{>0}$. A d -horizon optimal sequence $\mathbf{u}_d^*(x)$ such that $V_d(x) = J_{\gamma, d}(x, \mathbf{u}_d^*(x))$ exists for any $d \in \mathbb{Z}_{>0}$ in view of (3.3) as the input set \mathcal{U} is finite. Let $Y_d := V_d + W$, where W comes from item (ii) of SA2. We have that: (a) $\underline{\alpha}_Y(\sigma(x)) \leq Y_d(x) \leq \bar{\alpha}_Y(\sigma(x))$ holds with $\underline{\alpha}_Y = \alpha_W$ and $\bar{\alpha}_Y = \bar{\alpha}_V + \bar{\alpha}_W$ as [21, Theorem

1] or [24, Theorem 1] applies; (b) with $Y_0(\phi(d, x, \mathbf{u}_d^*(x))) \leq (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^{(d)}(Y_d(x))$ with $\alpha_Y := \alpha_W$ according to [21, (32)], and $\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1} \in \mathcal{K}_\infty$ (see footnote †2 in page 71). By applying (b) in (a), we obtain $\sigma(\phi(d, x, \mathbf{u}_d^*(x))) \leq \underline{\alpha}_Y^{-1} \circ (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^d(\bar{\alpha}_Y(\sigma(x)))$, and item (i) of Lemma B.1 holds. In view of $(\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}) \in \mathcal{K}_\infty$, it follows that $s - \alpha_Y \circ \bar{\alpha}_Y^{-1}(s) < s$ for any $s > 0$. Hence, by composing $(\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})$ to both sides $d-1$ times, we conclude $(\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^{(d)}(s) < (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^{(d-1)}(s)$ when $s > 0$, i.e. strictly decreasing in d for $s > 0$, and 0 when $s = 0$. Item (ii) of Lemma B.1 holds.

We now prove item (iii) of Lemma B.1. We have $\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1} \in \mathcal{K}_\infty$ such that $(\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})(s) < s$ for $s > 0$. For any $s \geq 0$, it follows that there exist $\varphi(s)$ such that $\lim_{k \rightarrow \infty} (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^{(k)}(s) = \varphi(s)$ and $\varphi(s) \leq (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^{(k)}(s)$ for all $k \in \mathbb{Z}_{\geq 0}$. We proceed by contradiction. Suppose there exist $s > 0$ such that $\varphi(s) \neq 0$. It follows that^{†1} $(\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})(\varphi(s)) = \varphi(s)$. Hence, $\varphi(s) = (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})(\varphi(s)) < \varphi(s)$, a contradiction is attained. Thus $\varphi(s) = 0$ for any $s \geq 0$. In other words, statement of item (iii) is verified. ■

B.2 Semiglobal asymptotic stability property for $\text{OP}_{\min, \text{stop}}$

The next result ensures a semiglobal asymptotic stability property for $\text{OP}_{\min, \text{stop}}$ and system (3.17).

Corollary B.1. *Suppose the following holds.*

(i) *There exist $L, \bar{a}_V, a_W > 0$, such that SA2 holds with $\bar{\alpha}_V(s) \leq \bar{a}_V s$, $\bar{\alpha}_W(s) \leq \bar{a}_W s$ and $\alpha_W(s) \geq a_W s$ for any $s \in [0, L]$.*

(ii) *For any $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}^{n_\varepsilon}$, $c_{\text{stop}}(\varepsilon, x) \leq |\varepsilon| \sigma(x)$.*

^{†1} Since $\lim_{k \rightarrow \infty} (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^{(k)}(s) = \varphi(s)$ implies $(\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})(\lim_{k \rightarrow \infty} (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^{(k)}(s)) = (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})(\varphi(s))$ as $\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1}$ is continuous, we have $(\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})(\varphi(s)) = \lim_{k \rightarrow \infty} (\mathbb{I} - \alpha_Y \circ \bar{\alpha}_Y^{-1})^{(k+1)}(\varphi(s)) = \varphi(s)$.

Let $\varepsilon^* > 0$ and $\Delta > L$ be such that

$$\varepsilon^* < \min \left\{ 1, \frac{a_W}{\bar{a}_V}, \frac{\bar{\alpha}_V^{-1}(\frac{1}{2}\alpha_W(L))}{\Delta} \right\}. \quad (\text{B.1})$$

Then, there exists $\beta \in \mathcal{KL}$, such that, for any $x \in \{z \in \mathbb{R}^n : \sigma(z) \leq \Delta\}$ and $\varepsilon \in \mathbb{R}^{n_\varepsilon}$ such that $|\varepsilon| < \varepsilon^*$, any solution $\phi(\cdot, x)$ to system (3.17) satisfies

$$\sigma(\phi(k, x)) \leq \beta(\sigma(x), k) \quad (\text{B.2})$$

for all $k \in \mathbb{Z}_{\geq 0}$. □

Proof. Let $\Delta > 0$, $x \in \mathbb{R}^n$ be such that $\sigma(x) \leq \Delta$. We select ε^* as in (B.1) and let $\varepsilon \in \mathbb{R}^{n_\varepsilon}$ such that $|\varepsilon| \leq \varepsilon^*$ and $v \in F_\varepsilon^*(x)$. Note that ε^* in (B.1) is well defined since the right-hand side is strictly positive. We will follow the same arguments as proof of Theorem 3.4, however applying the sublinear bounds of Corollary B.1. From item (ii) of Proposition 3.5, $Y_\infty(v) - Y_\infty(x) \leq -\alpha_W(\sigma(x)) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x))$. We use the following strategy. First, we show that $Y_\infty(v) - Y_\infty(x) \leq -\frac{\mu}{\bar{a}_V + \bar{a}_W} Y_\infty(x)$ holds for some $\mu > 0$ when $\sigma(x) \in [0, L]$ since $\varepsilon^* < \frac{a_W}{\bar{a}_V}$. Then, we show that $Y_\infty(v) - Y_\infty(x) \leq -\frac{1}{2}\tilde{\alpha}_Y(Y_\infty(x))$ holds for $\tilde{\alpha}_Y = \alpha_W \circ \bar{\alpha}_V^{-1}$ when $\sigma(x) \in (L, \Delta]$, given $\varepsilon^* < \frac{\bar{\alpha}_V^{-1}(\frac{1}{2}\alpha_W(L))}{\Delta}$. To conclude, we combine the two inequalities and we defer to the proof of Theorem 3.4.

Let, for the moment, x be such that $\sigma(x) \leq L$. From item (i) of Corollary B.1, we have that $-\alpha_W(\sigma(x)) \leq -a_W\sigma(x)$ holds as $\sigma(x) \leq L$, and similarly that $\bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x)) \leq \bar{a}_V|\varepsilon|\sigma(x)$, since $|\varepsilon|L < \varepsilon^*L < L$ follows from item (ii) of Corollary B.1 and (B.1). It follows then that $Y_\infty(v) - Y_\infty(x) \leq -\alpha_W(\sigma(x)) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x)) \leq (-a_W + \bar{a}_V|\varepsilon|)\sigma(x)$ holds. Since $|\varepsilon| \leq \varepsilon^* < \frac{a_W}{\bar{a}_V}$, we derive $-a_W + \bar{a}_V|\varepsilon| < 0$, hence, there exists $\mu > 0$ such that $-a_W + \bar{a}_V|\varepsilon| < -\mu$. We derive $Y_\infty(v) - Y_\infty(x) \leq -\mu\sigma(x)$. On the other hand, we have $Y_\infty(x) \leq \bar{\alpha}_Y(\sigma(x)) \leq (\bar{a}_V + \bar{a}_W)\sigma(x)$ in view of item (i) of Proposition 3.5, hence, $-(\bar{a}_V + \bar{a}_W)\sigma(x) \leq -Y_\infty(x)$. Since $\mu > 0$, we derive

$$Y_\infty(v) - Y_\infty(x) \leq -\frac{\mu}{\bar{a}_V + \bar{a}_W} Y_\infty(x). \quad (\text{B.3})$$

When $\sigma(x) \in (L, \Delta]$, it follows that $Y_\infty(v) - Y_\infty(x) \leq -\alpha_W(\sigma(x)) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, x)) \leq -\alpha_W(\sigma(x)) + \bar{\alpha}_V(\varepsilon^*\Delta)$ from item (ii) of Corollary B.1. As $\varepsilon^* < \frac{\bar{\alpha}_V^{-1}(\frac{1}{2}\alpha_W(L))}{\Delta}$, $\bar{\alpha}_V(\varepsilon^*\Delta) < \frac{1}{2}\alpha_W(L)$ holds. Since $\sigma(x) > L$ and $\alpha_W \in \mathcal{K}_\infty$, we have that $\frac{1}{2}\alpha_W(L) < \frac{1}{2}\alpha_W(\sigma(x))$, hence

$\bar{\alpha}_V(\varepsilon^* \Delta) < \frac{1}{2} \alpha_W(\sigma(x))$ and $Y_\infty(v) - Y_\infty(x) < -\alpha_W(\sigma(x)) + \frac{1}{2} \alpha_W(\sigma(x)) = -\frac{1}{2} \alpha_W(\sigma(x))$. Then, in view of item (i) of Proposition 3.5, we have $Y_\infty(x) \leq \bar{\alpha}_Y(\sigma(x))$ that implies $\alpha_W \circ \bar{\alpha}_Y^{-1}(Y_\infty(x)) \leq \alpha_W(\sigma(x))$, and conclude

$$Y_\infty(v) - Y_\infty(x) \leq -\frac{1}{2} \tilde{\alpha}_Y(Y_\infty(x)), \quad (\text{B.4})$$

where $\tilde{\alpha}_Y = \alpha_W \circ \bar{\alpha}_Y^{-1}$.

We have found that $Y_\infty(v) - Y_\infty(x)$ decreases for all $Y_\infty(x) \in (0, \bar{\alpha}_Y(\Delta)]$. In particular, by $-\frac{\mu}{\bar{a}_V + \bar{a}_W} Y_\infty(x)$ for $\sigma(x) \in [0, L]$ and by $-\frac{1}{2} \tilde{\alpha}_Y(Y_\infty(x))$ elsewhere, that is $Y_\infty(v) - Y_\infty(x) \leq -\min \left\{ \frac{\mu}{\bar{a}_V + \bar{a}_W} \mathbb{I}, \frac{1}{2} \tilde{\alpha}_Y \right\} (Y_\infty(x))$. The desired result is then derived by following the final steps of Theorem 3.4 to construct β . \blacksquare

The stability property in (B.2) corresponds to (3.18) with $\delta = 0$, thus ensuring a semiglobal asymptotic stability property. Inequality (B.1) can always be verified by taking ε^* small, since the right-hand side is strictly positive.

B.3 Average cost and cost-to-attractor guarantees for

OP_{min,stop}

In general, Theorem 3.4 only guarantess practical stability, thus the running cost $\mathcal{V}_\varepsilon^{\text{run}}(x)$ is not guarantee to be finite. We thus look at the average cost defined as

$$\mathcal{V}_\varepsilon^{\text{avg}}(x) := \left\{ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^N \ell_{U_\varepsilon^*(\phi(k,x))}(\phi(k,x)) : \phi(\cdot, x) \text{ is a solution to (3.17)} \right\}. \quad (\text{B.5})$$

As before, $\mathcal{V}_\varepsilon^{\text{avg}}(x)$ is

a set of possible averages, where $V_\varepsilon^{\text{avg}}(x) \in \mathcal{V}_\varepsilon^{\text{run}}(x)$ is the average of a possible solution of (3.17). We provide the next guarantee on each element of $\mathcal{V}_\varepsilon^{\text{avg}}(x)$

Theorem B.2. *Consider system (3.17), and suppose Assumption 3.1 and Theorem 3.4 hold with tuple $(\varepsilon^*, \delta, \Delta)$. For any ε such that $|\varepsilon| < \varepsilon^*$, any $x \in \{z \in \mathbb{R}^n : \sigma(z) \leq \Delta\}$, and $V_\varepsilon^{\text{avg}}(x) \in \mathcal{V}_\varepsilon^{\text{avg}}(x)$,*

$$0 \leq V_\varepsilon^{\text{avg}}(x) \leq \bar{\alpha}_V(\theta(|\varepsilon|, \delta)), \quad (\text{B.6})$$

where $\bar{\alpha}_V$ and θ comes from SA2 and Assumption 3.1, respectively. \square

Proof. Let $\Delta, \delta > 0$, $x \in \mathbb{R}^n$ such that $\sigma(x) \leq \Delta$. We select ε^* as in Theorem 3.4. Let $\varepsilon \in \mathbb{R}^{n_\varepsilon}$ such that $|\varepsilon| < \varepsilon^*$, $\phi(k+1, x) \in F_\varepsilon^*(\phi(k, x))$ for any $k \in \mathbb{Z}_{\geq 0}$ where ϕ is a solution to (3.17) initialized at x . For the sake of convenience, we denote $\ell(x, u) := \ell_u(x)$ for any $u \in \mathcal{U}$. Consider

$$V_\varepsilon^{\text{avg}}(x) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^N \ell(\phi(k, x), u_k^r), \quad (\text{B.7})$$

where $u_k^r \in \mathcal{U}_\varepsilon^*(\phi(k, x))$ such that $\phi(k+1, x) = f_{u_k^r}(\phi(k, x))$. Note that indeed $V_\varepsilon^{\text{avg}}(x) \in \mathcal{V}_\varepsilon^{\text{cta}}(x)$. The lower-bound $0 \leq V_\varepsilon^{\text{avg}}(x)$ in (B.6) follows immediately from $\ell(x, u) \geq 0$ for any $x \in \mathbb{R}^n$ and $u \in \mathcal{U}$. On the other hand, we derive from (3.23) that, for any $k \in \mathbb{Z}_{\geq 0}$,

$$\begin{aligned} & V_\infty(\phi(k+1, x)) - V_\varepsilon(\phi(k, x)) \\ & \leq -\ell(\phi(k, x), u_k^r) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, \phi(k, x))). \end{aligned} \quad (\text{B.8})$$

Hence

$$\ell(\phi(k, x), u_k^r) \leq V_\varepsilon(\phi(k, x)) - V_\infty(\phi(k+1, x)) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, \phi(k, x))), \quad (\text{B.9})$$

from which we deduce, for any $N \geq 0$,

$$\begin{aligned} \sum_{k=0}^N \ell(\phi(k, x), u_k^r) & \leq V_\varepsilon(\phi(0, x)) - V_\infty(\phi(1, x)) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, \phi(0, x))) \\ & \quad + V_\varepsilon(\phi(1, x)) - V_\infty(\phi(2, x)) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, \phi(1, x))) \\ & \quad + \dots \\ & \quad + V_\varepsilon(\phi(N(x), x)) - V_\infty(\phi(N+1, x)) + \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, \phi(N(x), x))) \\ & \leq V_\varepsilon(\phi(0, x)) + \sum_{k=0}^N \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, \phi(k, x))), \end{aligned} \quad (\text{B.10})$$

since $V_\varepsilon(\phi(k, x)) - V_\infty(\phi(k, x)) \leq 0$ for all $k \in \mathbb{Z}_{\geq 0}$ according to Theorem 2.3 and $V_\infty(\phi(N+1, x)) \geq 0$. According to Assumption 3.1, $c_{\text{stop}}(\varepsilon, \phi(k, x)) \leq \theta(|\varepsilon|, \sigma(\phi(k, x)))$, and since Theorem 3.4 holds, $\sigma(\phi(k, x)) \leq \max\{\beta(k, \sigma(x)), \delta\}$ as $\sigma(x) \leq \Delta$. Hence, by direct substitution in (B.7),

$$V_\varepsilon^{\text{avg}}(x) \leq \lim_{N \rightarrow \infty} \frac{1}{N} \left(V_\varepsilon(\phi(0, x)) + \sum_{k=0}^N \bar{\alpha}_V(\theta(\varepsilon, \max\{\beta(k, \sigma(x)), \delta\})) \right).$$

We break the sum in two parts. Let $H(x) \in \mathbb{Z}$ be such that $\beta(\sigma(x), k) > \delta$ for $k \in \{0, \dots, H(x)\}$ and $\beta(\sigma(x), k) \leq \delta$ for $k \in \{H(x) + 1, \dots\}$ when $\sigma(x) > \delta$, otherwise, if $\sigma(x) \leq \delta$, we define it $H(x) = -1$. Integer $H(x)$ exists and is finite since $\beta \in \mathcal{KL}$. It follows that $\beta(\sigma(x), k) \leq \delta$ for $k \in \{H(x) + 1, \dots\}$. Hence

$$V_\varepsilon^{\text{avg}}(x) \leq \lim_{N \rightarrow \infty} \frac{1}{N} \left(V_\varepsilon(x) + \sum_{k=0}^{\min\{H(x), N\}} \bar{\alpha}_V(\theta(\varepsilon, \beta(k, \sigma(x)))) + \sum_{k=H(x)+1}^N \bar{\alpha}_V(\theta(\varepsilon, \delta)) \right), \quad (\text{B.11})$$

where $\sum_{k=0}^{-1} = 0$ by convention. It follows that $\frac{1}{N} \left(V_\varepsilon(x) + \sum_{k=0}^{\min\{H(x), N\}} \bar{\alpha}_V(\theta(\varepsilon, \beta(k, \sigma(x)))) \right) \rightarrow 0$ as $N \rightarrow \infty$, which implies

$$V_\varepsilon^{\text{avg}}(x) \leq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=H(x)+1}^N \bar{\alpha}_V(\theta(|\varepsilon|, \delta)) \leq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^N \bar{\alpha}_V(\theta(|\varepsilon|, \delta)). \quad (\text{B.12})$$

Hence $V_\varepsilon^{\text{avg}}(x) \leq \bar{\alpha}_V(\theta(|\varepsilon|, \delta)) \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^N 1$ and Theorem B.2 holds. \blacksquare

Theorem B.2 shows that if δ is small, so is the average running cost and the latter can be made as close to 0 as desired by taking $|\varepsilon|$ small, as $\theta(\cdot, \delta) \in \mathcal{K}$ according to Assumption 3.1 and $\bar{\alpha}_V \in \mathcal{K}_\infty$. Note that the average cost associated to the *infinite*-horizon cost (3.2) is zero as $V_\infty(x) < \infty$ according to SA1. Hence, the mismatch between the latter and the elements of $\mathcal{V}_\varepsilon^{\text{avg}}(x)$ can be made as small as desired. Furthermore, the upper-bound in (B.5) is uniform with respect to x . Theorem 3.4 plays a vital role in Theorem B.2, as it guarantees that the state converges to the attractor $\{z \in \mathbb{R}^n : \sigma(z) \leq \delta\}$ from any initial condition $x \in \{z \in \mathbb{R}^n : \sigma(z) \leq \Delta\}$.

The average cost provides information about the performance along solutions to (3.17) at the long run, typically once these have converged to attractor $\{x \in \mathbb{R}^n : \sigma(x) \leq \delta\}$. To quantify performance in the ‘transient’, i.e. before the solution has entered and stays forever in the attractor, we propose to consider what we call the ‘cost-to-attractor’ function defined as

$$\mathcal{V}_\varepsilon^{\text{cta}}(x) := \left\{ \sum_{k=0}^{N(x)} \ell_{\mathcal{U}_\varepsilon^*(\phi(k, x))}(\phi(k, x)) : \phi(\cdot, x) \text{ is a solution to (3.17)} \right\}, \quad (\text{B.13})$$

where $N(x) > 0$ is an integer such that for every $n > N(x)$, $\sigma(\phi(n, x)) \leq \delta$. That is, in contrast to $V_\varepsilon^{\text{run}}(x)$ in (3.39), where the series goes up to infinity, here we truncate

the series ‘earlier’ at $N(x)$, when the state has reached once and for all the attractor $\{z \in \mathbb{R}^n : \sigma(z) \leq \delta\}$. Given the semiglobal practical stability property of Theorem 3.4, $N(x)$ is well-defined by selecting ε^* sufficiently small. We give the following property for $\mathcal{V}_\varepsilon^{cta}(x)$.

Theorem B.3. *Consider system (3.17), and suppose Assumption 3.1 and Theorem 3.4 hold with tuple $(\varepsilon^*, \delta, \Delta)$. For any ε such that $|\varepsilon| < \varepsilon^*$, any $x \in \{z \in \mathbb{R}^n : \sigma(z) \in (\delta, \Delta]\}$, and $V_\varepsilon^{cta}(x) \in \mathcal{V}_\varepsilon^{cta}(x)$,*

$$0 \leq V_\varepsilon^{cta}(x) \leq V_\varepsilon(x) + \sum_{k=0}^{N(x)} \bar{\alpha}_V(\theta(|\varepsilon|, \max\{\beta(\sigma(x), k), \delta\})), \quad (\text{B.14})$$

where $\bar{\alpha}_V$, β and θ comes from SA2, Theorem 3.4 and Assumption 3.1, respectively. \square

Proof. Let $\Delta, \delta > 0$, $x \in \mathbb{R}^n$ such that $\sigma(x) \leq \Delta$. We select ε^* as in Theorem 3.4. Let $\varepsilon \in \mathbb{R}^{n_\varepsilon}$ such that $|\varepsilon| < \varepsilon^*$, $\phi(k+1, x) \in F_\varepsilon^*(\phi(k, x))$ for any $k \in \mathbb{Z}_{\geq 0}$ where ϕ is a solution to (3.17) initialized at x , and $N(x)$ is such that for any $n > N(x)$, $\phi(n, x) \leq \delta$ which exists since $\sigma(\phi(k, x)) \leq \max\{\beta(\sigma(x), k), \delta\}$ for any $k \in \mathbb{Z}_{\geq 0}$ and $\beta \in \mathcal{KL}$ according to (3.18). For the sake of convenience, we denote $\ell(x, u) = \ell_u(x)$ as in the proof of Theorem B.2. Consider

$$V_\varepsilon^{cta}(x) := \sum_{k=0}^{N(x)} \ell(\phi(k, x), u_k^r), \quad (\text{B.15})$$

where $u_k^r \in \mathcal{U}_\varepsilon^*(\phi(k, x))$ such that $\phi(k+1, x) = f_{u_k^r}(\phi(k, x))$. From (B.10) in the proof of Theorem B.2, we have that

$$\sum_{k=0}^N \ell(\phi(k, x), u_k^r) \leq V_\varepsilon(\phi(0, x)) + \sum_{k=0}^N \bar{\alpha}_V(c_{\text{stop}}(\varepsilon, \phi(k, x))), \quad (\text{B.16})$$

for any $N \geq 0$. Theorem B.3 holds by taking $N = N(x)$, and by invoking Assumption 3.1 and Theorem 3.4, that is, $\bar{\alpha}_V(c_{\text{stop}}(\varepsilon, \phi(k, x))) \leq \bar{\alpha}_V(\theta(|\varepsilon|, \max\{\beta(\sigma(x), k), \delta\}))$. \blacksquare

Theorem B.3 implies that the ‘cost-to-attractor’ $\{x \in \mathbb{R}^n : \sigma(x) \leq \delta\}$ at x is upper-bounded by V_ε and an error term which can be controlled by ε . In contrast to the average cost in Theorem B.2, we observe in Theorem B.3 a role of the decay rate of β on the cost-to-attractor: the faster $\beta(\cdot, k)$ decays, the smaller $N(x)$ and the smaller the error term in (B.13).

Appendix C. Additional simulation plots for Chapter 3

What follows is a sequence of plots for $OP_{\min, \text{stop}}$ for system (3.43) for various ε and initial state $x = (10, 15)$. We can think of $\frac{1}{\varepsilon}$ as a ‘gain’, that affects the stability of the closed loop system from unstable (Figure C.1), to marginally stable (Figure C.2), to ‘first order’ responses with progressively faster decay rates (Figures C.3,C.4,C.5), to a final ‘second order’ response (Figure C.6). This is expected by Corollary 3.2, as (3.35) predicts faster decay rates for smaller $|\varepsilon|$. We note that no better sequence was found for $\varepsilon \leq \frac{1}{9} = 0.111$.

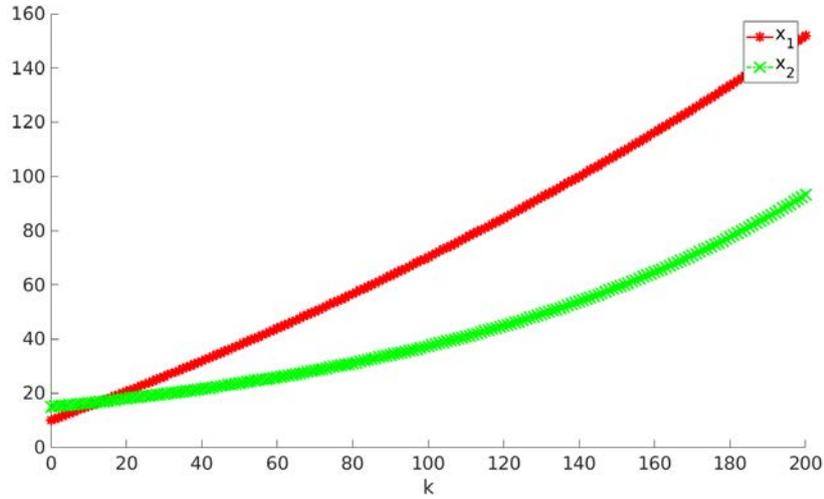


Figure C.1: State evolution for $OP_{\min, \text{stop}}$ with $\varepsilon = 5$ and $x = (10, 15)$.

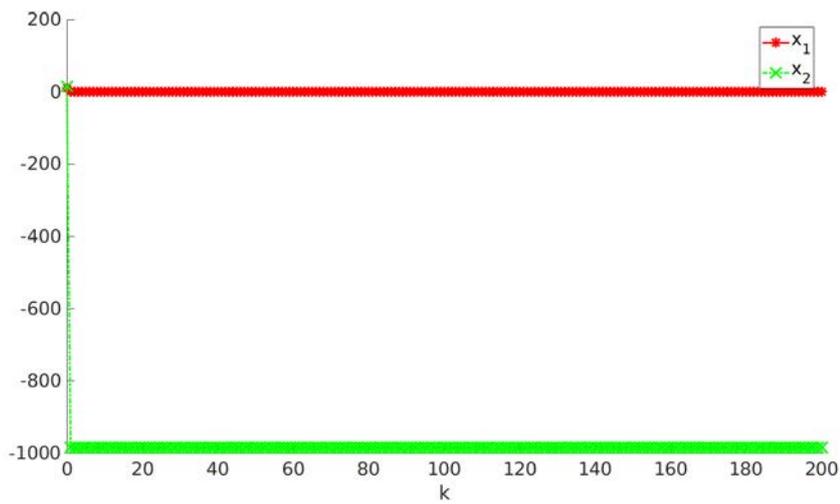


Figure C.2: State evolution for $OP_{\min, \text{stop}}$ with $\varepsilon = \frac{10}{9} = 1.111$ and $x = (10, 15)$.

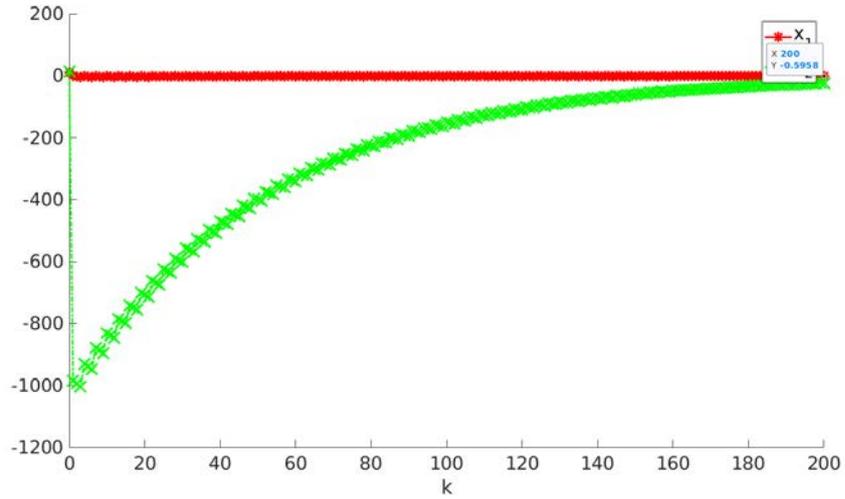


Figure C.3: State evolution for $OP_{\min, \text{stop}}$ with $\varepsilon = \frac{10}{11} = 0.909$ and $x = (10, 15)$.

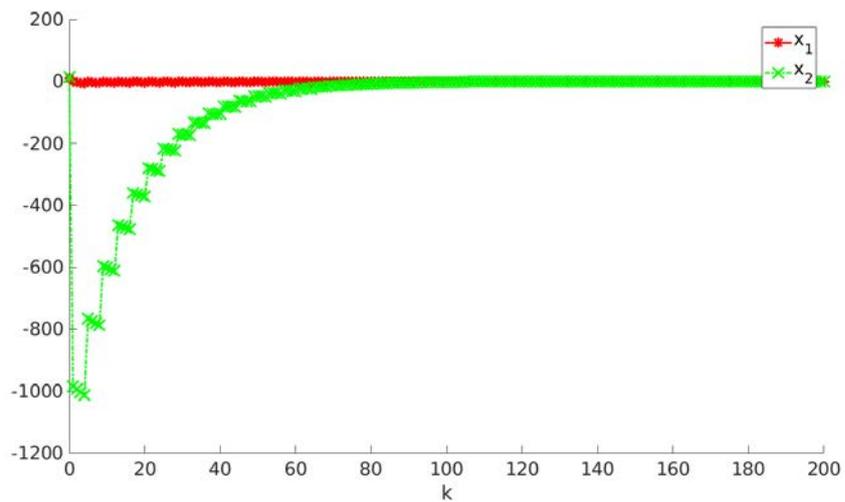


Figure C.4: State evolution for $OP_{\min, \text{stop}}$ with $\varepsilon = \frac{10}{12} = 0.833$ and $x = (10, 15)$.

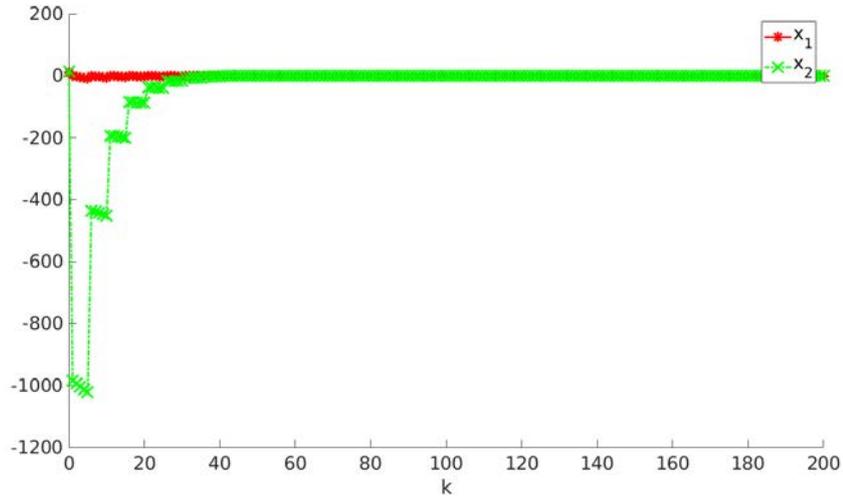


Figure C.5: State evolution for $OP_{\min, \text{stop}}$ with $\varepsilon = \frac{10}{17} = 0.588$ and $x = (10, 15)$.

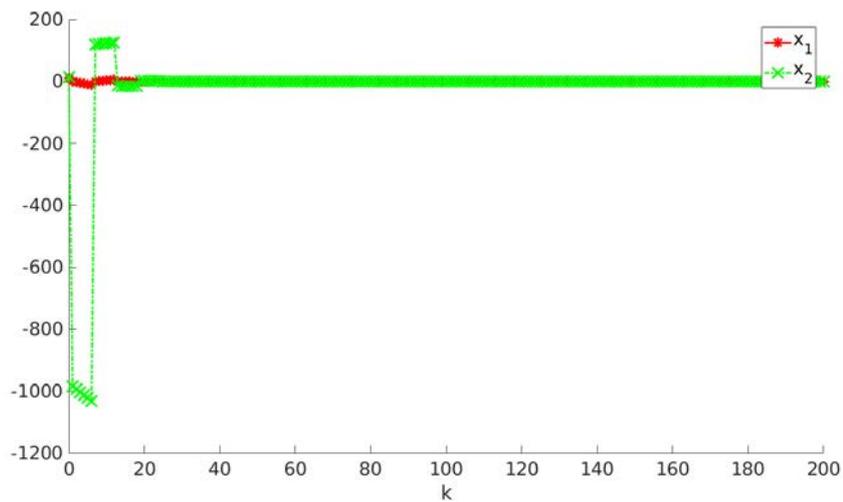


Figure C.6: State evolution for $OP_{\min, \text{stop}}$ with $\varepsilon = \frac{1}{9} = 0.111$ and $x = (10, 15)$.

Appendix D. Stability guarantees for nonlinear discrete-time systems controlled by approximate value iteration

Parallel to the work presented in this manuscript, this PhD candidate also collaborated on the following related conference paper [56].

Stability guarantees for nonlinear discrete-time systems controlled by approximate value iteration

Romain Postoyan, Mathieu Granzotto, Lucian Buşoniu, Bruno Scherrer, Dragan Nešić and Jamal Daafouz

Abstract— Value iteration is a method to generate optimal control inputs for generic nonlinear systems and cost functions. Its implementation typically leads to approximation errors, which may have a major impact on the closed-loop system performance. We talk in this case of approximate value iteration (AVI). In this paper, we investigate the stability of systems for which the inputs are obtained by AVI. We consider deterministic discrete-time nonlinear plants and a class of general, possibly discounted, costs. We model the closed-loop system as a family of systems parameterized by tunable parameters, which are used for the approximation of the value function at different iterations, the discount factor and the iteration step at which we stop running the algorithm. It is shown, under natural stabilizability and detectability properties as well as mild conditions on the approximation errors, that the family of closed-loop systems exhibit local practical stability properties. The analysis is based on the construction of a Lyapunov function given by the sum of the approximate value function and the Lyapunov-like function that characterizes the detectability of the system. By strengthening our conditions, asymptotic and exponential stability properties are guaranteed.

I. INTRODUCTION

Value iteration (VI) is one of the pillars of dynamic programming, which allows generating optimal control inputs for general nonlinear systems and cost functions. While optimality is the primary concern in the dynamic programming literature, recent results in e.g. [11], [22], [27] provide conditions under which the input sequence generated by VI stabilizes the origin of the closed-loop system. Nevertheless, a major downside of VI is its computational complexity, which makes it intractable in general. To overcome this issue, the so-called approximate value iteration (AVI) was proposed, leading to sub-optimal policies due to the induced approximation errors [1], [3]. These errors must be carefully handled as they may have a major impact on the system performance. Thus, an important problem in approximate dynamic programming is the analysis of the impact of the

approximation errors on the obtained value function, see e.g., [1], [10], [16], [18], [19], [23], [24]. In this context, a discount factor is often introduced in the cost function to ensure that the value function does not blow up when iterating AVI [1].

While the impact of approximation errors on optimality is largely covered in the literature, e.g., [1], [3], [6], [10], [18], [19], their effect on the stability properties of the closed-loop system is less understood. The authors of [4] proposed a model-free version of AVI for linear stochastic and deterministic continuous-time systems with quadratic costs for which near-optimality and stability are ensured. In [17], [26], nonlinear discrete-time systems with generic positive definite stage costs are investigated. Asymptotic stability properties are ensured provided conditions involving the optimal value function are satisfied, but these conditions may be difficult to verify. To overcome this potential issue, the author of [12] provides explicit conditions on the approximation errors under which local asymptotic stability of the origin is guaranteed. The assumptions made in [12] may be restrictive though. First, the approximation errors are required to vanish in the attractor, which may be difficult to ensure. Second, the proved stability property is asymptotic, while we know, from the parallel between VI and model-predictive control [2] that, in general, only practical stability can be achieved for nonlinear systems [9]. Third, common points of [4], [12], [17], [26] are that the stability of the origin is studied, while the closed-loop system may have a more general type of attractor. Finally, the considered cost functions are undiscounted in these references, while discounted costs are customary in dynamic programming. We aim at relaxing these limitations in this paper.

We consider nonlinear deterministic discrete-time systems and general non-negative cost functions, possibly discounted. Stability is investigated in terms of a generic measuring function, thus covering the stability of the origin and of more general compact sets in a unified way, see e.g., [7], [9], [14], [21], [25]. This allows addressing situations in which the closed-loop system exhibits a limit cycle for instance. Approximation errors are considered in the evaluation of the approximate value function, in the approximate feedback policy and in the criterion at which we stop iterating the algorithm. These errors are parameterized by a vector of tunable parameters denoted ε , which model the meta-parameters, such as the number of points in an interpolation grid or the number of neurons in a neural network, used in the employed approximation scheme. We model the overall system as a family of systems parameterized by ε , a possible discount

R. Postoyan, M. Granzotto and J. Daafouz are with the Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France (e-mails: firstname.name@univ-lorraine.fr).

L. Buşoniu is with the Department of Automation, Technical University of Cluj-Napoca, Memorandumului 28, 400114 Cluj-Napoca, Romania (e-mail: lucian.busoniu@aut.utcluj.ro). His work was supported by a grant of the Romanian Ministry of Research and Innovation, CNCS - UEFISCDI, project number PN-III-P1-1.1-PD-2016-1304, contract PD 27/2018, within PNCDI III.

B. Scherrer is with the Université de Lorraine, CNRS, Inria, IECL, F-54000 Nancy, France (e-mail: bruno.scherrer@univ-lorraine.fr).

D. Nešić is with the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, VIC 3010, Australia (e-mail: dnesic@unimelb.edu.au). His work was supported by the Australian Research Council under the Discovery Project DP170104099.

factor and the iteration step at which the algorithm is stopped. We make stabilizability and detectability assumptions on the plant and the stage cost like in [7], [9], [21]. We then impose conditions on the approximation errors, which state that these errors can be made as small as desired by suitably tuning ε (i.e. by increasing the computational power used when doing the approximations). We do not require the errors to vanish in the attractor as in [12], and these conditions do not depend on the optimal value function as in [17], [26]. Based on these assumptions and inspired by [9], [21], we construct a Lyapunov-like function given by the sum of the approximate value function and a function used in the detectability assumption. We then ensure a local practical stability property. Asymptotic and exponential stability properties are also derived by strengthening the assumptions. We finally discuss the relationship between our results and [12, Theorem 2].

Compared to [8, Section V], where stability results applicable to AVI are provided, the analysis in this paper does not rely on the knowledge of near-optimality bounds but on the properties of the algorithm itself. This provides a clearer delineation of the conditions under which the algorithm will provide appropriate stability properties. This also allows us to rely on more natural assumptions. As a result, we are able to ensure an asymptotic stability property, as opposed to a practical one, in the absence of approximation errors and when there is no discount factor, contrary to [8].

The rest of the paper is organized as follows. The problem is formally stated in Section II. The assumptions are given in Section III. The main results are presented in Section IV. Section V provides conclusions. Long proofs have been postponed to the appendix for the sake of readability.

Notation. Let \mathbb{R} be the set of real numbers, $\mathbb{R}_{\geq 0} := [0, \infty)$, $\mathbb{R}_{> 0} := (0, \infty)$, \mathbb{Z} be the set of integers, $\mathbb{Z}_{\geq 0} := \{0, 1, 2, \dots\}$ and $\mathbb{Z}_{> 0} := \{1, 2, \dots\}$. We denote by $\mathcal{B}_n(\varepsilon)$ the closed ball of \mathbb{R}^n centered at the origin of radius $\varepsilon > 0$, $n \in \mathbb{Z}_{> 0}$. The notation (x, y) stands for $[x^\top, y^\top]^\top$, where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. A function $\chi : [0, a) \rightarrow \mathbb{R}_{\geq 0}$ with $a \in \mathbb{R}_{> 0} \cup \{\infty\}$ is of class \mathcal{K} if it is continuous, zero at zero and strictly increasing, and it is of class \mathcal{K}_∞ if, in addition, $a = \infty$ and it is unbounded. A continuous function $\chi : [0, a) \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ with $a \in \mathbb{R}_{> 0} \cup \{\infty\}$ is of class \mathcal{KL} if for each $t \in \mathbb{R}_{\geq 0}$, $\chi(\cdot, t)$ is of class \mathcal{K} , and, for each $s \in [0, a)$, $\chi(s, \cdot)$ is decreasing and converges to zero at infinity. In the particular case where $\chi : (s_1, s_2) \mapsto \lambda_1 s_1 e^{-\lambda_2 s_2}$ for some $\lambda_1 \in [1, \infty)$ and $\lambda_2 > 0$, we write that $\chi \in \exp -\mathcal{KL}$. A function $\chi : [0, a) \rightarrow \mathbb{R}_{\geq 0}$, with $a \in \mathbb{R}_{> 0} \cup \{\infty\}$, is of class \mathcal{N} when it is non-decreasing. We say that a function $\chi : [0, a_1) \times [0, a_2) \rightarrow \mathbb{R}_{\geq 0}$, with $a_1, a_2 \in \mathbb{R}_{> 0} \cup \{\infty\}$, is of class \mathcal{KN} when $\chi(\cdot, s)$ is of class \mathcal{K} for any $s \in [0, a_2)$ and $\chi(s, \cdot)$ is of class \mathcal{N} for any $s \in [0, a_1)$. The notation \mathbb{I} stands for the identity map from $\mathbb{R}_{\geq 0}$ to $\mathbb{R}_{\geq 0}$. The Euclidean norm of vector $x \in \mathbb{R}^n$ is denoted by $|x|$ and the distance of $x \in \mathbb{R}^n$ to a set $\mathcal{A} \subseteq \mathbb{R}^n$ is denoted by $|x|_{\mathcal{A}} := \inf\{|x - y| : y \in \mathcal{A}\}$.

II. PROBLEM STATEMENT

Consider the nonlinear discrete-time system

$$x_{k+1} = f(x_k, u_k), \quad (1)$$

where $x_k \in \mathbb{R}^{n_x}$ is the state, $u_k \in \mathcal{U}(x_k)$ is the control input, $\mathcal{U}(x_k) \subseteq \mathbb{R}^{n_u}$ is the set of admissible inputs for state x_k , $k \in \mathbb{Z}_{\geq 0}$ is the time, $f : \mathcal{W} \rightarrow \mathbb{R}^{n_x}$ with $\mathcal{W} := \{(x, u) : x \in \mathbb{R}^{n_x}, u \in \mathcal{U}(x)\}$, and $n_x, n_u \in \mathbb{Z}_{> 0}$.

The cost function is given by

$$J_\gamma(x, \mathbf{u}) := \sum_{k=0}^{\infty} \gamma^k \ell(\phi(k, x, \mathbf{u}|_k), u_k), \quad (2)$$

where $x \in \mathbb{R}^{n_x}$, $\mathbf{u} = (u_0, u_1, \dots)$ is an infinite-length sequence of admissible inputs, $\gamma \in (0, 1]$ is possibly sub-unitary, $\ell : \mathcal{W} \rightarrow \mathbb{R}_{\geq 0}$ is the stage cost, which takes non-negative values, and $\phi(k, x, \mathbf{u}|_k)$ is the solution to (1) at the k^{th} -step starting at state x with input sequence $\mathbf{u}|_k := (u_0, \dots, u_{k-1})$, which is the truncation of \mathbf{u} to the first $k \in \mathbb{Z}_{> 0}$ steps¹.

We investigate the scenario where approximate value iteration [1] is employed to approximately minimize cost (2) over the sequence of inputs. We thus start from an initial value function $\hat{V}^0 : \mathbb{R}^{n_x} \rightarrow \mathbb{R}_{\geq 0}$. We then iterate it as follows, for any $x \in \mathbb{R}^{n_x}$ and $i \in \mathbb{Z}_{\geq 0}$,

$$\hat{V}_{\gamma, \varepsilon_V}^{i+1}(x) = \min_{u \in \mathcal{U}(x)} \left[\ell(x, u) + \gamma \hat{V}_{\gamma, \varepsilon_V}^i(f(x, u)) \right] + e_{\hat{V}, i+1}(\gamma, \varepsilon_V, x), \quad (3)$$

where $\hat{V}_{\gamma, \varepsilon_V}^0 := \hat{V}^0$ and $e_{\hat{V}, i+1} : (0, 1] \times \mathcal{B}_{n_V}(\bar{\varepsilon}) \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}$ is an approximation error function. In general, there are two sources of error. The first is due to the limited approximation power of the given function space in which $\hat{V}_{\gamma, \varepsilon_V}^i$ is required to lie, characterized e.g. by the so-called inherent Bellman error in [19], or by the smallest distance between any point in the function space and the optimal value function, see e.g. [5]. Depending on the particular approximation scheme employed, and if the value function satisfies certain smoothness properties, it may be possible to control this error arbitrarily close to zero by increasing the approximation power of the function space, which is done by tuning meta-parameters such as the number of points on an interpolation grid, number of neurons in a neural network etc. This is related to the so-called universal function approximation property that many approximators have been proved to satisfy. The variation of the error with these meta-parameters is represented by the variation of $\varepsilon_V \in \mathbb{R}^{n_V}$. This vector of parameters is constrained in $\mathcal{B}_{n_V}(\bar{\varepsilon})$, where $\bar{\varepsilon} \in \mathbb{R}_{> 0}$ is an upper-bound on the norm of ε_V , and $n_V \in \mathbb{Z}_{> 0}$. Additionally, there is often a second source of error due to the limited number of samples used to compute $\hat{V}_{\gamma, \varepsilon_V}^i$. The guarantees given in [19] for stochastic problems ensure that

¹We use the convention $\phi(0, x, \mathbf{u}|_0) = x$ where $\mathbf{u}|_0$ is the empty set.

²The adjustable parameter ε_V , later ε , is constrained in $\mathcal{B}_{n_V}(\bar{\varepsilon})$, later $\mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$. The forthcoming results straightforwardly apply when we have additional requirements, by constraining accordingly $\varepsilon_V, \varepsilon$, in the forthcoming statements.

this error goes to arbitrarily close to zero with probability arbitrarily close to 1, if the number of samples is large enough. Sometimes, e.g. in the deterministic case, when the approximator is interpolative and the samples are equal to the grid centers, this source of error can be removed [5].

We make the next standing assumption.

Standing Assumption 1 (SA1): For any $\gamma \in (0, 1]$, $\varepsilon_V \in \mathcal{B}_{n_V}(\bar{\varepsilon})$, $x \in \mathbb{R}^{n_x}$ and $i \in \mathbb{Z}_{\geq 0}$, $\hat{V}_{\gamma, \varepsilon_V}^i(x) \geq 0$. \square

SA1 is natural in the context of the paper as the objective of (3) is to iteratively estimate the optimal value function associated to (2), which is non-negative as $\ell(\mathcal{W}) \subseteq \mathbb{R}_{\geq 0}$. It can be enforced by suitably selecting the function space where $\hat{V}_{\gamma, \varepsilon_V}^i$, $i \in \mathbb{Z}_{\geq 0}$, is constrained to lie.

The policy, or feedback, calculated at iteration $i \in \mathbb{Z}_{\geq 0}$, is given by

$$\hat{h}_{\gamma, \varepsilon_h}^i(x) \in \operatorname{argmin}_{u \in \mathcal{U}(x)} \left[\ell(x, u) + \gamma \hat{V}_{\gamma, \varepsilon_V}^i(f(x, u)) \right] + e_{\hat{h}, i}(\gamma, \varepsilon_h, x), \quad (4)$$

where $e_{\hat{h}, i} : (0, 1] \times \mathcal{B}_{n_h}(\bar{\varepsilon}) \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_u}$ is another approximation error function, which may arise when computing the argmin above, and which can be tuned via parameter $\varepsilon_h \in \mathcal{B}_{n_h}(\bar{\varepsilon})$, $n_h \in \mathbb{Z}_{>0}$. The norm of ε_h is subject to the same constraint as ε_V , namely $|\varepsilon_h| \leq \bar{\varepsilon}$, without loss of generality as we can always define $\bar{\varepsilon}$ as the minimum of the respective bounds on the norms of ε_h and ε_V . Error $e_{\hat{h}, i}$ must be such that the next standing assumption holds.

Standing Assumption 2 (SA2): For any $\gamma \in (0, 1]$, $\varepsilon_h \in \mathcal{B}_{n_h}(\bar{\varepsilon})$, $x \in \mathbb{R}^{n_x}$ and $i \in \mathbb{Z}_{\geq 0}$, $\hat{h}_{\gamma, \varepsilon_h}^i(x) \in \mathcal{U}(x)$. \square

We stop the iterating procedure (3) at any step i bigger than $i^* \in \mathbb{Z}_{>0}$, which is assumed to satisfy the next property.

Standing Assumption 3 (SA3): There exist $i^* \in \mathbb{Z}_{\geq 0}$ and $e_{\text{stop}, i} : (0, 1] \times \mathcal{B}_{n_V}(\bar{\varepsilon}) \times \mathcal{B}_{n_{\text{stop}}}(\bar{\varepsilon}) \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}_{\geq 0}$ with $n_{\text{stop}} \in \mathbb{Z}_{>0}$, such that for any $\gamma \in (0, 1]$, $\varepsilon_V \in \mathcal{B}_{n_V}(\bar{\varepsilon})$, $\varepsilon_{\text{stop}} \in \mathcal{B}_{n_{\text{stop}}}(\bar{\varepsilon})$, $x \in \mathbb{R}^{n_x}$ and $i \geq i^*$,

$$\mathcal{T}\hat{V}_{\gamma, \varepsilon_V}^i(x) \leq \hat{V}_{\gamma, \varepsilon_V}^i(x) + e_{\text{stop}, i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) \quad (5)$$

where $\mathcal{T}\hat{V}_{\gamma, \varepsilon_V}^i(x) := \min_{u \in \mathcal{U}(x)} \left[\ell(x, u) + \gamma \hat{V}_{\gamma, \varepsilon_V}^i(f(x, u)) \right]$. \square

Error $e_{\text{stop}, i}$ is called the Bellman residual [28] of $\hat{V}_{\gamma, \varepsilon_V}^i$, $i \geq i^*$. It is allowed to depend on γ , ε_V , x and also $\varepsilon_{\text{stop}} \in \mathcal{B}_{n_{\text{stop}}}(\bar{\varepsilon})$ a vector of additional adjustable parameters. The expression of $e_{\text{stop}, i}$ allows both the usual iteration error like in (3), and an additional error due e.g. to stopping the algorithm early (i.e. before some asymptotic regime is exactly reached). The latter component typically arises implicitly as a result of the stopping condition of the algorithm.

Remark 1: It is important to note that typical bounds in the literature on performance guarantees are given on the distance between the optimal value function and the last value function iterate $\hat{V}_{\gamma, \varepsilon_V}^i$. It can then easily be shown that the Bellman residual in (5) is at most $(1 + \gamma)|\hat{V}_{\gamma, \varepsilon_V}^i(x) - V_{\gamma, \mathcal{X}}^*(x)|$, where $V_{\gamma, \mathcal{X}}^*$ is the optimal value function of cost (2). \square

We concatenate the adjustable parameters arising in the approximation errors as $\varepsilon := (\varepsilon_V, \varepsilon_h, \varepsilon_{\text{stop}}) \in \mathbb{R}^{n_\varepsilon}$ where $n_\varepsilon := n_h + n_V + n_{\text{stop}}$. We write in the following the

approximate value function as $\hat{V}_{\gamma, \varepsilon}^i$ instead of $\hat{V}_{\gamma, \varepsilon_V}^i$ for the sake of convenience. We similarly write $\hat{h}_{\gamma, \varepsilon}^i$ instead of $\hat{h}_{\gamma, \varepsilon_h}^i$.

As a result, the closed-loop feedback system is given by

$$\begin{aligned} x_{k+1} &\in f\left(x_k, \operatorname{argmin}_{u_k \in \mathcal{U}(x_k)} \left[\ell(x_k, u_k) + \gamma \hat{V}_{\gamma, \varepsilon}^i(f(x_k, u_k)) \right] \right. \\ &\quad \left. + e_{\hat{h}, i^*}(\gamma, \varepsilon_h, x_k) \right) \\ &=: \hat{F}_{\gamma, \varepsilon, i}(x_k). \end{aligned} \quad (6)$$

Equation (6) describes a family of systems parameterized by γ , ε and i . Note that (6) indeed depends on ε_V and $\varepsilon_{\text{stop}}$ through (3) and (5).

Our objective is to analyse stability properties of the family of systems (6). For this purpose, we make assumptions on the detectability of system (1) with respect to the stage cost ℓ , on the approximation errors $e_{\hat{V}, i}$, $e_{\hat{h}, i}$ and $e_{\text{stop}, i}$, and on the stabilizability of system (1).

III. ASSUMPTIONS

We use a generic measuring function of the state $\sigma : \mathbb{R}^{n_x} \rightarrow \mathbb{R}_{\geq 0}$ to investigate stability like in [9], [21], which is required to satisfy the next assumption.

Assumption 1: Function σ is continuous on \mathbb{R}^{n_x} .

Assumption 1 is verified when σ is defined as $|\cdot|$, $|\cdot|^2$ or $x \mapsto x^\top P x$ with P a real, symmetric and positive definite matrix, when studying the stability of $x = 0$. This condition is also verified when σ is given by $|\cdot|_{\mathcal{A}}$ or $|\cdot|_{\mathcal{A}}^2$ with $\mathcal{A} \subseteq \mathbb{R}^n$, when studying the stability of set \mathcal{A} .

A. Detectability

We make the next detectability assumption on system (1) with respect to ℓ as in [7], [9], [21].

Assumption 2: There exist a continuous function $W : \mathbb{R}^{n_x} \rightarrow \mathbb{R}_{\geq 0}$, $\alpha_W \in \mathcal{K}_\infty$ and $\bar{\alpha}_W : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ continuous, nondecreasing and zero at zero, such that the following holds for any $(x, u) \in \mathcal{W}$

$$\begin{aligned} W(x) &\leq \bar{\alpha}_W(\sigma(x)) \\ W(f(x, u)) - W(x) &\leq -\alpha_W(\sigma(x)) + \ell(x, u). \end{aligned} \quad (7)$$

\square

Assumption 2 is indeed a detectability property. The best way to see it is when (7) holds with $W = 0$. Then, (7) implies that $\ell(x, u) \geq \alpha_W(\sigma(x))$, which means that, if $\ell(x, u) = 0$ (is small), then $\sigma(x) = 0$ (is small) as $\alpha_W \in \mathcal{K}_\infty$. Note that Assumption 2 is independent of γ and the approximation errors as it only involves stage cost ℓ and system (1). Assumption 2 generalizes the common requirement that $\ell(x, u)$ is positive definite as in e.g., [12], [17], [22], [27].

Remark 2: It is possible to relax Assumption 2 by replacing the second inequality in (7) with $W(f(x, u)) - W(x) \leq -\alpha_W(\sigma(x)) + \chi(\ell(x, u))$, where $\chi \in \mathcal{K}_\infty$, as in [9, SA3]. This extension is left for future work. \square

B. Approximation errors

We make assumptions on the approximation errors $e_{\hat{V},i}$, $e_{\hat{h},i}$ and $e_{\text{stop},i}$ with $i \geq i^*$ in order to prove stability properties for system (6). For this purpose, we constrain the state to a given compact set $\mathcal{X} \subseteq \mathbb{R}^{n_x}$. This is justified by the fact that AVI is often applied on a compact subset of the state space, which corresponds to \mathcal{X} here. We make the next assumption on \mathcal{X} .

Assumption 3: The set $\{x \in \mathbb{R}^n : \sigma(x) = 0\}$ is included in the interior of \mathcal{X} . \square

Assumption 3 is verified when σ is defined as the distance to a compact set for example, and this set is included in the interior of \mathcal{X} .

For any $\gamma \in (0, 1]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, $x \in \mathcal{X}$, consider $\hat{h}_{\gamma,\varepsilon}^i(x)$ as in (4). We define

$$h_{\gamma,\varepsilon}^i(x) := \hat{h}_{\gamma,\varepsilon}^i(x) - e_{\hat{h},i}(\gamma, \varepsilon_h, x). \quad (8)$$

which corresponds to the feedback $\hat{h}_{\gamma,\varepsilon}^i$ in (4) not affected by the error term $e_{\hat{h},i}$. Hence $h_{\gamma,\varepsilon}^i(x) \in \operatorname{argmin}_{u \in \mathcal{U}(x)} [\ell(x, u) + \gamma \hat{V}_{\gamma,\varepsilon}^i(f(x, u))]$. We make the next assumption.

Assumption 4: There exist $\alpha_{\text{stop}}, \hat{\alpha}_V, \hat{\alpha}_\ell \in \mathcal{KN}$, $\underline{\alpha}_W, \underline{\alpha}_\varepsilon \in \mathcal{K}$ such that, for any $\gamma \in (0, 1]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, $x \in \mathcal{X}$ and $i \geq i^*$,

$$e_{\text{stop},i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) \leq \alpha_{\text{stop}}(|\varepsilon|, \sigma(x)) \quad (9a)$$

$$\min \left\{ |e_{\hat{V},i}(\gamma, \varepsilon_V, x)|, e_{\text{stop},i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) \right\} \leq \alpha_W(\sigma(x)) - \underline{\alpha}_W(\sigma(x)) + \underline{\alpha}_\varepsilon(|\varepsilon|) \quad (9b)$$

$$\left| \hat{V}_{\gamma,\varepsilon}^i(f(x, h_{\gamma,\varepsilon}^i(x))) - \hat{V}_{\gamma,\varepsilon}^i(f(x, \hat{h}_{\gamma,\varepsilon}^i(x))) \right| \leq \hat{\alpha}_V(|\varepsilon|, \sigma(x)) \quad (9c)$$

$$\left| \ell(x, h_{\gamma,\varepsilon}^i(x)) - \ell(x, \hat{h}_{\gamma,\varepsilon}^i(x)) \right| \leq \hat{\alpha}_\ell(|\varepsilon|, \sigma(x)). \quad (9d)$$

\square

Condition (9a) implies that the iteration (3) practically converges, so that the mismatch $\mathcal{T}\hat{V}_{\gamma,\varepsilon}^i - \hat{V}_{\gamma,\varepsilon}^i$ can be made as small as desired by reducing $|\varepsilon|$. This condition covers properties previously assumed in the literature. For instance, [4, Algorithm 1] ensures (9a) by taking $\alpha_{\text{stop}}(|\varepsilon|, \sigma(x)) = |\varepsilon| \lambda_{\max}(P_i) \sigma(x)$ where $\lambda_{\max}(P_i)$ is the maximum eigenvalue of symmetric, positive definite matrix P_i , $\varepsilon = \varepsilon_i \bar{\varepsilon}$ and $\sigma(x) = |x|^2$, using the notation of this reference. Also, [17, (3.2.10)] implies that (9a) holds with $\alpha_{\text{stop}}(|\varepsilon|, \sigma(x)) = |\varepsilon|$. In this case, that is when (9a) is satisfied with α_{stop} independent of $\sigma(x)$, (9b) follows with $\underline{\alpha}_W = \alpha_W$ and $\underline{\alpha}_\varepsilon = \alpha_{\text{stop}}$. Similarly, when $|e_{\hat{V},i}(\gamma, \varepsilon_V, x)| \leq \bar{\alpha}_{V,\varepsilon}(|\varepsilon|)$ with $\bar{\alpha}_{V,\varepsilon} \in \mathcal{K}$ for any $x \in \mathcal{X}$ and $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, (9b) is verified with $\underline{\alpha}_W = \alpha_W$ and $\underline{\alpha}_\varepsilon = \bar{\alpha}_{V,\varepsilon}$. In this case, Lemma 1 in [19] provides conditions under which (9a) holds with α_{stop} , which only depends on ε .

The last two conditions of Assumption 4 state that the mismatch between the values of $\ell(x, \cdot)$ and $\hat{V}_{\gamma,\varepsilon}^i(f(x, \cdot))$ for $i \geq i^*$ evaluated at the applied policy $\hat{h}_{\gamma,\varepsilon}^i(x)$ and at the ideal one $h_{\gamma,\varepsilon}^i(x)$ can be made as small as desired by tuning

ε . The next lemma provides sufficient conditions to ensure (9c)-(9d). Its proof is omitted for space reasons.

Lemma 1: Suppose the following holds.

- (i) For any $i \geq i^*$, $\hat{V}_{\gamma,\varepsilon}^i$ is continuous on \mathcal{X} .
- (ii) There exists $\alpha_\sigma \in \mathcal{K}_\infty$ such that $\sigma(x) \geq \alpha_\sigma(|x|)$ for any $x \in \mathbb{R}^n$.
- (iii) One of the next conditions holds.
 - (iii-a) There exists $\theta_\ell \in \mathcal{K}$ such that $|\ell(x, u) - \ell(x, v)| \leq \theta_\ell(|u - v|)$ for any $x \in \mathcal{X}$ and $(u, v) \in \mathcal{U}(x)^2$.
 - (iii-b) There exist $\alpha_\sigma \in \mathcal{K}_\infty$ and $\vartheta_h : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ non-decreasing such that, for any $\gamma \in (0, 1]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, $i \geq i^*$, $x \in \mathcal{X}$, $\sigma(x) \geq \alpha_\sigma(|x|)$ and $|h_{\gamma,\varepsilon}^i(x)| \leq \vartheta_h(\sigma(x))$.
- (iii) There exist $\alpha_h \in \mathcal{KN}$ such that, for any $\gamma \in (0, 1]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, $i \geq i^*$ and $x \in \mathcal{X}$, $|e_{\hat{h},i}(\gamma, \varepsilon_h, x)| \leq \alpha_h(|\varepsilon_h|, \sigma(x))$.

Then (9c)-(9d) hold. \square

The continuity of $\hat{V}_{\gamma,\varepsilon}^i$ in item (i) of Lemma 1 can be enforced by suitably selecting the function space where $\hat{V}_{\gamma,\varepsilon}^i$ is constrained to lie. Item (ii) is verified when σ is defined as the distance to a compact set, which includes the origin, for example. Item (iii-a) of Lemma 1 is verified when the stage cost can be written as $\ell(x, u) = \ell_1(x) + \ell_2(u)$ for any $x \in \mathcal{X}$ and $u \in \mathcal{U}(x) = \mathcal{V}$ where ℓ_2 is continuous and \mathcal{V} is compact for instance. In this case, ℓ_2 is uniformly continuous on \mathcal{V} by Heine theorem and ϑ_ℓ corresponds to the modulus of continuity, see Proposition A.2.1 in [20]. Regarding item (iii-b) of Lemma 1, the condition on σ is verified for the examples provided at the beginning of this section for example when the set \mathcal{A} is compact. The condition on $h_{\gamma,\varepsilon}^i$, on the other hand, is verified when $\mathcal{U}(x) \subseteq \mathcal{V}$ with \mathcal{V} compact to give an example, as it suffices to define ϑ_h as the constant function whose value is the maximum norm of the elements of \mathcal{V} . Finally, item (iv) implies that the error $e_{\hat{h},i}(\gamma, \varepsilon_h, x)$ in (4) can be made as small as desired by tuning ε_h .

C. Stabilizability

We make the next assumption on the approximate value function $\hat{V}_{\gamma,\varepsilon}^i$ with $i \geq i^*$.

Assumption 5: There exist $\bar{\alpha}_V, \bar{\alpha}_\varepsilon \in \mathcal{K}_\infty$ such that for any $\gamma \in (0, 1]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, $i \geq i^*$, $x \in \mathcal{X}$,

$$\hat{V}_{\gamma,\varepsilon}^i(x) \leq \bar{\alpha}_V(\sigma(x)) + \bar{\alpha}_\varepsilon(|\varepsilon|). \quad (10)$$

\square

Assumption 5 is related to the stabilizability of system (1). To see it, assume there exists $\check{\alpha}_V \in \mathcal{KN}$ such that for any $\gamma \in (0, 1]$, $\varepsilon_V \in \mathcal{B}_{n_V}(\bar{\varepsilon})$, $x \in \mathcal{X}$ and $i \in \mathbb{Z}_{\geq 0}$, $|e_{\hat{V},i}(\gamma, \varepsilon_V, x)| \leq \check{\alpha}_V(|\varepsilon_V|, \sigma(x))$. Such an assumption is generally needed for the satisfaction of the first two inequalities in (9). Define the modified stage cost $\tilde{\ell}(x, u) := \ell(x, u) + \check{\alpha}_V(\bar{\varepsilon}, \sigma(x))$ for any $(x, u) \in \mathcal{W}$. The associated *undiscounted* optimal value

function is $\tilde{V}(x) := \min_{u \in \mathcal{U}(x)} \sum_{k=0}^{\infty} \tilde{\ell}(\phi(k, x, \mathbf{u}|_k), u_k)$ for any

$x \in \mathcal{X}$. The next lemma gives a condition on \tilde{V} ensuring Assumption 5.

Lemma 2: Under Assumption 4, if there exists $\tilde{\alpha}_V \in \mathcal{K}_\infty$ such that, for any $x \in \mathcal{X}$,

$$\tilde{V}(x) \leq \tilde{\alpha}_V(\sigma(x)) \quad (11)$$

then Assumption 5 holds with $\bar{\alpha}_V = \tilde{\alpha}_V$ and $\bar{\alpha}_\varepsilon = 0$. \square

Proof. The proof is similar to [10, Section V.B]. In view of the definition of $\tilde{\ell}$ and Assumption 4, $\gamma\ell(x, u) \leq \ell(x, u) \leq \tilde{\ell}(x, u)$ for any $(x, u) \in \mathcal{W}$ and $\gamma \in (0, 1]$. Hence, by definition of \tilde{V} , for any $\gamma \in (0, 1]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, $x \in \mathcal{X}$, $i \geq i^*$, it holds $\hat{V}_{\gamma, \varepsilon}^i(x) \leq \tilde{V}(x)$. The desired result is obtained by invoking (11). \blacksquare

Condition (11) is a stabilizability property of system (1) with respect to stage cost $\tilde{\ell}$ according to [9, Section III] and in [21, Lemma 1], which we recall below.

Lemma 3: Consider system (1) and suppose that $\tilde{\ell}$ is *uniformly globally exponentially stabilizable to zero with respect to* $\sigma : \mathbb{R}^{n_x} \rightarrow \mathbb{R}_{\geq 0}$, i.e. there exist $M > 0$ and $\lambda > 0$, where λ is called the *decrease rate*, such that for any $x \in \mathbb{R}^{n_x}$ there exists an admissible infinite-length control input sequence $\mathbf{u}(x)$ verifying $\tilde{\ell}(\phi(k, x, \mathbf{u}|_k(x)), u_k(x)) \leq M\sigma(x)e^{-\lambda k}$ for any $k \in \mathbb{Z}_{>0}$. Then (11) holds with $\tilde{\alpha}_V(s) = \frac{Ms}{1-e^{-\lambda}}$ for any $s \geq 0$. \square

The works in [17], [26] require (10) to hold with $\bar{\alpha}_\varepsilon = 0$, see the proof of Theorem 4 in [26] for instance. On the other hand, condition (11) typically requires $e_{\hat{V}, i+1}(\gamma, \varepsilon_V, x)$ to vanish in the attractor $\{x \in \mathcal{X} : \sigma(x) = 0\}$ to hold, which may be difficult to ensure. When this is not the case, Assumption 5 is more likely to be verified with $\bar{\alpha}_\varepsilon \neq 0$ in (10).

Remark 3: As mentioned in the introduction, [8, Section V] also allows investigating stability of nonlinear discrete-time systems controlled by AVI. While Assumption 2 is also considered in [8], Assumptions 4 and 5 are more closely related to the approximation errors, which we can tune when implementing AVI, and not on near-optimality bounds as in [8, Assumption 3] that may be subject to some conservatism. \square

IV. STABILITY GUARANTEES

A. Main result

Based on Assumptions 1-5, we can state the next theorem about the existence of a Lyapunov-like function for system (6).

Theorem 1: Under Assumptions 1-5, there exist $\alpha_Y, \underline{\alpha}_Y, \bar{\alpha}_Y \in \mathcal{K}_\infty$ and $\Upsilon \in \mathcal{KN}$ such that for any $(\gamma, \varepsilon) \in (0, 1] \times \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, and $i \geq i^*$ where i^* comes from SA3, function $Y_{\gamma, \varepsilon}^i := \hat{V}_{\gamma, \varepsilon}^i + W$ defined on \mathbb{R}^{n_x} satisfies the following.

- For any $x \in \mathcal{X}$, $\underline{\alpha}_Y(\sigma(x)) - \underline{\alpha}_\varepsilon(|\varepsilon|) \leq Y_{\gamma, \varepsilon}^i(x) \leq \bar{\alpha}_Y(\sigma(x)) + \bar{\alpha}_\varepsilon(|\varepsilon|)$.
- For any $x \in \mathcal{X}$, $v \in \hat{F}_{\gamma, \varepsilon, i}(x)$, $Y_{\gamma, \varepsilon}^i(v) - Y_{\gamma, \varepsilon}^i(x) \leq -\alpha_Y(\sigma(x)) + \Upsilon\left(\frac{1-\gamma}{\gamma} + |\varepsilon|, \sigma(x)\right)$, where $\hat{F}_{\gamma, \varepsilon, i}$ is defined in (6).

The expressions of $\underline{\alpha}_Y$, $\bar{\alpha}_Y$, α_Y and Υ are provided in Table I. \square

We are ready to state the main stability result.

Theorem 2: Consider system (6) and suppose Assumptions 1-5 hold. There exists $\beta \in \mathcal{KL}$ such that for any $\delta \in \mathbb{R}_{>0}$, there exists $(\gamma^*, \epsilon^*, \Delta) \in (0, 1) \times (0, \bar{\varepsilon}) \times \mathbb{R}_{>0}$ such that for any $i \geq i^*$, $\gamma \in [\gamma^*, 1]$, $\epsilon \in [0, \epsilon^*]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, $x \in \mathcal{X}$ with $Y_{\gamma, \varepsilon}^i(x) \leq \Delta$, where $Y_{\gamma, \varepsilon}^i$ is defined in Theorem 1, any solution³ $\phi(\cdot, x)$ to system (6) satisfies $\phi(k, x) \in \mathcal{X}$ and

$$\sigma(\phi(k, x)) \leq \beta(\sigma(x), k) + \delta, \quad (12)$$

for all $k \in \mathbb{Z}_{\geq 0}$. \square

Theorem 2 ensures a local practical stability property with respect to σ , in the sense that given *any* neighborhood of the set $\{x \in \mathcal{X} : \sigma(x) = 0\}$, there exists a tuple composed of a lower bound on the discount factor γ , an upper-bound on the norm of the tunable parameters ε and an upper-bound Δ on the initial value of the Lyapunov function such that, for any pair (γ, ε) satisfying this bound, any solution initialized at x such that $Y_{\gamma, \varepsilon}^i(x) \leq \Delta$ lies in \mathcal{X} for all positive times and converges to the desired neighborhood of $\{x \in \mathcal{X} : \sigma(x) = 0\}$. It is possible to ensure stronger properties by strengthening the conditions of Theorem 2, see Section IV-B.

Theorem 2, as well as the stability results presented in the sequel, can in principle be used to compute bounds on the minimum allowable discount factor γ^* , the maximal error ϵ^* , see (28), and Δ . However, first, it is likely that the obtained values are subject to some conservatism because of the used proof techniques. Second, tight bounds would also require precise characterizations of the functions introduced in Assumptions 2, 4 and 5, which may be difficult to achieve. In this case, our stability results can be applied in a qualitative way, meaning that for large enough γ and small enough ϵ and Δ , stability follows and the type of stability depends on the type of functions in Assumptions 2, 4 and 5, and not their precise expressions.

Remark 4: In the absence of errors, i.e. $\varepsilon = 0$, and when $\gamma = 1$, the set $\{x : \sigma(x) = 0\}$ is locally asymptotically stable. This follows directly from the proof of Theorem 2. This result is consistent with [11, Theorem 1] where σ is the Euclidean distance, which is generalized here to a larger class of stage costs and to more general types of attractors. Moreover, this fact is an improvement over the general results in [8], which ensures a practical stability property in this case. \square

B. Stronger statements

The next result ensures an asymptotic and exponential stability properties for system (6) and the bounds γ^* and ϵ^* are independent of constant Δ , contrary to⁴ Theorem 2.

Corollary 1: Consider system (6) and suppose the following holds.

- Assumptions 1 and 3 hold.
- $e_{\hat{h}, i} = 0$ in (4) for any $i \in \mathbb{Z}_{\geq 0}$.

³We use the same notation to denote a solution to (6) and to (1), see (2), with some slight abuse.

⁴See (24).

$\underline{\alpha}_Y :=$	$\underline{\alpha}_W$
$\overline{\alpha}_Y :=$	$\overline{\alpha}_V + \overline{\alpha}_W$
$\alpha_Y :=$	α_W
$\Upsilon :=$	$(s_1, s_2) \mapsto \frac{1-\gamma}{\gamma} \left(\overline{\alpha}_V(s_2) + \overline{\alpha}_\varepsilon(s_1) + \alpha_{\text{stop}}(s_1, s_2) \right) + \widehat{\alpha}_\ell(s_1, s_2) + \alpha_{\text{stop}}(s_1, s_2) + \widehat{\alpha}_V(s_1, s_2)$

TABLE I
EXPRESSIONS OF THE FUNCTIONS USED IN THEOREM 1.

- (iii) Assumptions 2, 4 and 5 hold with $\alpha_W(s) = a_W \cdot s$, $\overline{\alpha}_V(s) = \overline{a}_V \cdot s$, $\overline{\alpha}_\varepsilon = 0$, $\alpha_{\text{stop}}(s_1, s_2) = a_{\text{stop}}(s_1)s_2$, $\underline{\alpha}_\varepsilon = 0$, where $a_W, \overline{a}_V > 0$ and $a_{\text{stop}} \in \mathcal{K}$, for any $s, s_1, s_2 \geq 0$.

Let $(\gamma^*, \epsilon^*) \in (0, 1) \times (0, \bar{\epsilon})$ be such that

$$\frac{1-\gamma^*}{\gamma^*} \left(\overline{a}_V + a_{\text{stop}}(\epsilon^*) \right) + a_{\text{stop}}(\epsilon^*) < a_W. \quad (13)$$

There exist $\beta \in \mathcal{KL}$ and $\Delta > 0$ such that for any $i \geq i^*$, $\gamma \in [\gamma^*, 1]$, $\epsilon \in [0, \epsilon^*]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\epsilon)$, and $x \in \mathcal{X}$ such that $Y_{\gamma, \varepsilon}^i(z) \leq \Delta$, any solution $\phi(\cdot, x)$ to system (6) initialized at $x \in \mathcal{X}$ satisfies $\phi(k, x) \in \mathcal{X}$ and

$$\sigma(\phi(k, x)) \leq \beta(\sigma(x), k), \quad (14)$$

for all $k \in \mathbb{Z}_{\geq 0}$. In addition, when $\underline{\alpha}_W(s) = \underline{a}_W \cdot s$ for any $s \geq 0$, $\beta \in \exp -\mathcal{KL}$. \square

Sufficient conditions that ensure the satisfaction of the properties required by Corollary 1 are given in the following. Corollary 1 differs from Theorem 2 on two points. First, property (14) is no longer practical but asymptotic (or exponential) as $\sigma(\phi(k, x))$ is guaranteed to converge to the origin as the time tends to infinity, contrary to (12). Second, the pair (γ^*, ϵ^*) is independent of Δ , contrary to Theorem 2. Note that condition (13) is always verified for γ^* sufficiently close to 1 and ϵ^* sufficiently small as $a_{\text{stop}} \in \mathcal{K}$.

To relate the conditions of Corollary 1 to those in [12, Theorem 2], we first need to adapt one of the assumptions in [12, Theorem 2]. Indeed, the work in [12] does not interpret system (1) controlled by AVI as a family of parameterized systems as we do, but as a single system with fixed parameters. However, if we view constant c in [12, Assumption 2] as a parameter, it would then correspond to $|\varepsilon_V|$ with our notation. We also need to add a few extra conditions, otherwise we do not see how to prove that solutions initialized in a neighborhood of the origin remains in the set \mathcal{X} , contrary to what [12, Theorem 2] states, as explained in the following. Once this is done, the obtained set of conditions imply the satisfaction of the requirements of Corollary 1. As a result, the origin of system (6) is locally asymptotically stable.

Corollary 2: Consider system (6) and suppose the following holds.

- (i) Set $\mathcal{X} \subset \mathbb{R}^{n_x}$ is compact, connected and the origin is in its interior.
- (ii) For any $(x, u) \in \mathcal{W}$, $\ell(x, u) = Q(x) + R(u)$ with Q and R continuous and positive definite.
- (iii) $\gamma = 1$ in (2).

- (iv) There exists $\bar{\epsilon} \in [0, 1)$ such that for any $\varepsilon_V \in \mathcal{B}_{n_V}(\bar{\epsilon})$ $|e_{\widehat{V}, i+1}(1, \varepsilon_V, x)| \leq |\varepsilon_V| \ell(x, 0)$ for any $x \in \mathcal{X}$.
- (v) For any $\varepsilon_V \in [0, \bar{\epsilon})$, $x \in \mathbb{R}^{n_x} \setminus \mathcal{X}$ and $i \in \mathbb{Z}_{\geq 0}$, $\widehat{V}_{\gamma, \varepsilon_V}^i(x) \geq 0$.
- (vi) For any $i \in \mathbb{Z}_{\geq 0}$, $\varepsilon_h \in \mathbb{R}^{n_h}$ and $x \in \mathbb{R}^{n_x}$, $e_{\widehat{h}, i}(1, \varepsilon_h, x) = 0$ in (4).
- (vii) There exists $d > 0$ such that $\widehat{V}^0(x) \leq dQ(x)$ for any $x \in \mathcal{X}$.

Then the conditions of Corollary 1 hold with $\sigma = Q$, $W = 0$, $\overline{\alpha}_W = \alpha_W = \mathbb{I}$, $a_W = 1$, $\overline{\alpha}_V = 2d\mathbb{I}$, $\overline{a}_V = 2d$, $\overline{\alpha}_\varepsilon = 0$, $\alpha_{\text{stop}}(s_1, s_2) = a_{\text{stop}}(s_1)d s_2$ with $a_{\text{stop}}(s_1) = \frac{4s_1}{1-s_1}$ for any $s_1 \in [0, 1)$ and $s_2 \geq 0$, $\underline{\alpha}_\varepsilon = 0$, $\underline{\alpha}_W = (1 - \bar{\epsilon})\mathbb{I}$ and $\underline{a}_W = 1 - \bar{\epsilon}$. \square

The conditions of Corollary 2 correspond to those in [12, Theorem 2] except that we have considered a parameterized version of [12, Assumption 2] in item (iv) of Corollary 2 as explained above, and we have added items (i) and (v). The latter are essential to prove the forward invariance of Lyapunov level set $\widehat{\mathcal{B}}_r^i$ as defined in [12, Theorem 2]. Indeed, the corresponding Lyapunov function needs to be non-negative on \mathbb{R}^{n_x} and not only on \mathcal{X} as written in [12], thus justifying item (v) of Corollary 2. Moreover, for set $\widehat{\mathcal{B}}_r^i$ to be not only the origin, the latter needs to be in the interior of \mathcal{X} as stated in item (i) of Corollary 2 (and \mathcal{X} does not actually need to be connected).

Corollary 2 indeed leads to a local asymptotic stability property for the origin of system (6) consistently with [12, Theorem 2], by application of Corollary 1. As Q is positive definite and continuous, there exist $\underline{\alpha}_Q, \overline{\alpha}_Q \in \mathcal{K}_\infty$ such that $\underline{\alpha}_Q(|x|) \leq Q(x) \leq \overline{\alpha}_Q(|x|)$ for any $x \in \mathcal{X}$, as \mathcal{X} is compact here, in view of⁵ [15, Lemma 4.3]. Hence, (14) leads to $|\phi(k, x)| \leq \tilde{\beta}(|x|, k)$ with $\tilde{\beta}(s_1, s_2) \mapsto \underline{\alpha}_Q^{-1}(\beta(\overline{\alpha}_Q(s_1), s_2)) \in \mathcal{KL}$, which is the standard characterization of (local) asymptotic stability of the origin. Moreover, condition (13) is, in this case, $\epsilon^* < \frac{1}{4d+1}$, which is less restrictive than the corresponding condition, namely $\epsilon^* < 1 + 2d - \sqrt{4d^2 + 4d}$, in [12, Theorem 2].

V. CONCLUSION

We have presented conditions under which a nonlinear discrete-time system whose inputs are generated by AVI satisfies stability properties. A key idea is to consider the closed-loop system as a family of systems parameterized by

⁵Lemma 4.3 in [15] ensures the existence of such functions $\underline{\alpha}_Q, \overline{\alpha}_Q \in \mathcal{K}$ but these functions can be taken in \mathcal{K}_∞ as \mathcal{X} is compact.

tunable algorithmic parameters, the discount factor and the iteration step at which we stop running the algorithm. Stability is defined using a generic measuring function, covering point and set stability. Stabilizability and detectability assumptions are made for this purpose, which are in line with conditions imposed in previous papers in different contexts [9], [21]. The approximation errors arising in the implementation of AVI are required to satisfy mild conditions. As a result, local practical stability properties are guaranteed, which become asymptotic or exponential under stronger conditions.

This work can be extended in various ways, among which the case where the policy is time-varying as in [12]. We will also investigate the issue of robustness, which is not trivial in discrete-time [14].

VI. APPENDIX

Proof of Theorem 1. Let $\gamma \in (0, 1]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, $i \geq i^*$, $x \in \mathcal{X}$ and $v \in \hat{F}_{\gamma, \varepsilon, i}(x)$. Hence, $v = f(x, \hat{h}_{\gamma, \varepsilon}^i(x))$ for some $\hat{h}_{\gamma, \varepsilon}^i(x)$ defined as in (4). We first prove item (i) of Theorem 1. In view of Assumptions 2 and 5 and the definition of $Y_{\gamma, \varepsilon}^i$, $Y_{\gamma, \varepsilon}^i(x) \leq \bar{\alpha}_V(\sigma(x)) + \bar{\alpha}_\varepsilon(|\varepsilon|) + \bar{\alpha}_W(\sigma(x))$. Hence $Y_{\gamma, \varepsilon}^i(x) \leq \bar{\alpha}_Y(\sigma(x)) + \bar{\alpha}_\varepsilon(|\varepsilon|)$ where $\bar{\alpha}_Y := \bar{\alpha}_V + \bar{\alpha}_W \in \mathcal{K}_\infty$. In view of SA3,

$$\hat{V}_{\gamma, \varepsilon}^i(x) \geq \ell(x, h_{\gamma, \varepsilon}^i(x)) + \gamma \hat{V}_{\gamma, \varepsilon}^i(f(x, h_{\gamma, \varepsilon}^i(x))) - e_{\text{stop}, i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x), \quad (15)$$

where $h_{\gamma, \varepsilon}^i(x)$ is as in (8). We deduce that $Y_{\gamma, \varepsilon}^i(x) = \hat{V}_{\gamma, \varepsilon}^i(x) + W(x) \geq \ell(x, h_{\gamma, \varepsilon}^i(x)) + \gamma \hat{V}_{\gamma, \varepsilon}^i(f(x, h_{\gamma, \varepsilon}^i(x))) - e_{\text{stop}, i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) + W(x)$. According to SA1, $\gamma \hat{V}_{\gamma, \varepsilon}^i(f(x, h_{\gamma, \varepsilon}^i(x))) \geq 0$, thus $Y_{\gamma, \varepsilon}^i(x) \geq \ell(x, h_{\gamma, \varepsilon}^i(x)) - e_{\text{stop}, i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) + W(x)$. On the other hand, according to Assumption 2, $W(x) \geq \alpha_W(\sigma(x)) - \ell(x, h_{\gamma, \varepsilon}^i(x))$. We derive $Y_{\gamma, \varepsilon}^i(x) \geq -e_{\text{stop}, i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) + \alpha_W(\sigma(x))$. In view of the last inequality in (9), we have, when $e_{\text{stop}, i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) \leq |e_{\hat{V}_{\gamma, \varepsilon}^i}(\gamma, \varepsilon_V, x)|$, $Y_{\gamma, \varepsilon}^i(x) \geq \underline{\alpha}_W(\sigma(x)) - \underline{\alpha}_\varepsilon(|\varepsilon|)$. If $e_{\text{stop}, i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) > |e_{\hat{V}_{\gamma, \varepsilon}^i}(\gamma, \varepsilon_V, x)|$, in view of (3),

$$\hat{V}_{\gamma, \varepsilon}^i(x) = \ell(x, h_{\gamma, \varepsilon}^{i-1}(x)) + \gamma \hat{V}_{\gamma, \varepsilon}^{i-1}(f(x, h_{\gamma, \varepsilon}^{i-1}(x))) + e_{\hat{V}_{\gamma, \varepsilon}^i}(\gamma, \varepsilon_V, x). \quad (16)$$

By following similar lines as above, we derive that $Y_{\gamma, \varepsilon}^i(x) \geq e_{\hat{V}_{\gamma, \varepsilon}^i}(\gamma, \varepsilon_V, x) + \alpha_W(\sigma(x)) \geq -|e_{\hat{V}_{\gamma, \varepsilon}^i}(\gamma, \varepsilon_V, x)| + \alpha_W(\sigma(x))$. Then, in view of the last inequality in (9), we derive $Y_{\gamma, \varepsilon}^i(x) \geq \underline{\alpha}_W(\sigma(x)) - \underline{\alpha}_\varepsilon(|\varepsilon|)$. Thus, item (i) of Theorem 1 holds with $\underline{\alpha}_Y = \underline{\alpha}_W \in \mathcal{K}_\infty$.

We now prove item (ii) of Theorem 1. We have that

$$Y_{\gamma, \varepsilon}^i(v) - Y_{\gamma, \varepsilon}^i(x) = \hat{V}_{\gamma, \varepsilon}^i(v) - \hat{V}_{\gamma, \varepsilon}^i(x) + W(v) - W(x). \quad (17)$$

In view of (15),

$$Y_{\gamma, \varepsilon}^i(v) - Y_{\gamma, \varepsilon}^i(x) \leq \hat{V}_{\gamma, \varepsilon}^i(v) - \ell(x, h_{\gamma, \varepsilon}^i(x)) - \gamma \hat{V}_{\gamma, \varepsilon}^i(f(x, h_{\gamma, \varepsilon}^i(x))) + e_{\text{stop}, i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) + W(v) - W(x). \quad (18)$$

We obtain, by adding and subtracting $\hat{V}_{\gamma, \varepsilon}^i(f(x, h_{\gamma, \varepsilon}^i(x)))$,

$$\begin{aligned} Y_{\gamma, \varepsilon}^i(v) - Y_{\gamma, \varepsilon}^i(x) &\leq \hat{V}_{\gamma, \varepsilon}^i(v) - \hat{V}_{\gamma, \varepsilon}^i(f(x, h_{\gamma, \varepsilon}^i(x))) \\ &\quad - \ell(x, h_{\gamma, \varepsilon}^i(x)) \\ &\quad + (1 - \gamma) \hat{V}_{\gamma, \varepsilon}^i(f(x, h_{\gamma, \varepsilon}^i(x))) \\ &\quad + e_{\text{stop}, i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) \\ &\quad + W(v) - W(x). \end{aligned} \quad (19)$$

From Assumption 2 and by definition of v , we have

$$W(v) - W(x) \leq -\alpha_W(\sigma(x)) + \ell(x, \hat{h}_{\gamma, \varepsilon}^i(x)). \quad (20)$$

Therefore, in view of (19),

$$\begin{aligned} Y_{\gamma, \varepsilon}^i(v) - Y_{\gamma, \varepsilon}^i(x) &\leq -\alpha_W(\sigma(x)) \\ &\quad + \ell(x, \hat{h}_{\gamma, \varepsilon}^i(x)) - \ell(x, h_{\gamma, \varepsilon}^i(x)) \\ &\quad + (1 - \gamma) \hat{V}_{\gamma, \varepsilon}^i(f(x, h_{\gamma, \varepsilon}^i(x))) \\ &\quad + e_{\text{stop}, i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) \\ &\quad + \hat{V}_{\gamma, \varepsilon}^i(v) - \hat{V}_{\gamma, \varepsilon}^i(f(x, h_{\gamma, \varepsilon}^i(x))). \end{aligned} \quad (21)$$

The rest of the proof consists in suitably upper-bounding $\ell(x, \hat{h}_{\gamma, \varepsilon}^i(x)) - \ell(x, h_{\gamma, \varepsilon}^i(x))$, $(1 - \gamma) \hat{V}_{\gamma, \varepsilon}^i(f(x, h_{\gamma, \varepsilon}^i(x)))$, $\hat{V}_{\gamma, \varepsilon}^i(v) - \hat{V}_{\gamma, \varepsilon}^i(f(x, h_{\gamma, \varepsilon}^i(x)))$ and $e_{\text{stop}, i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x)$ to obtain the desired result. In view of (15) and since ℓ takes non-negative values,

$$\hat{V}_{\gamma, \varepsilon}^i(f(x, h_{\gamma, \varepsilon}^i(x))) \leq \frac{1}{\gamma} \left(\hat{V}_{\gamma, \varepsilon}^i(x) + e_{\text{stop}, i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) \right). \quad (22)$$

Hence, in view of Assumptions 4 and 5,

$$\begin{aligned} (1 - \gamma) \hat{V}_{\gamma, \varepsilon}^i(f(x, h_{\gamma, \varepsilon}^i(x))) &\leq \frac{1 - \gamma}{\gamma} \left(\bar{\alpha}_V(\sigma(x)) + \bar{\alpha}_\varepsilon(|\varepsilon|) + \alpha_{\text{stop}}(|\varepsilon|, \sigma(x)) \right). \end{aligned} \quad (23)$$

Consequently, by applying Assumption 4, $Y_{\gamma, \varepsilon}^i(v) - Y_{\gamma, \varepsilon}^i(x) \leq -\alpha_W(\sigma(x)) + \hat{\alpha}_\ell(|\varepsilon|, \sigma(x)) + \frac{1 - \gamma}{\gamma} \left(\bar{\alpha}_V(\sigma(x)) + \bar{\alpha}_\varepsilon(|\varepsilon|) + \alpha_{\text{stop}}(|\varepsilon|, \sigma(x)) \right) + \alpha_{\text{stop}}(|\varepsilon|, \sigma(x)) + \hat{\alpha}_V(|\varepsilon|, \sigma(x)) = -\alpha_Y(\sigma(x)) + \Upsilon \left(\frac{1 - \gamma}{\gamma} + |\varepsilon|, \sigma(x) \right)$, with α_Y and Υ defined as in Table I, which are indeed of class \mathcal{K}_∞ and \mathcal{KN} , respectively, in view of the properties of the involved functions. We have proved that item (ii) of Theorem 1 holds. \blacksquare

Proof of Theorem 2. Let $\delta > 0$, $i \geq i^*$ and $\tilde{\varepsilon} \in (0, \bar{\varepsilon}]$ be small enough such that the set $\{x \in \mathbb{R}^{n_x} : \underline{\alpha}_Y(\sigma(x)) \leq \underline{\alpha}_\varepsilon(\tilde{\varepsilon})\}$ is included in the interior of \mathcal{X} , where $\underline{\alpha}_Y$ is defined in Table I and $\underline{\alpha}_\varepsilon$ comes from Assumption 4. Such a constant $\tilde{\varepsilon}$ always exists in view of Assumptions 1 and 3.

We introduce several quantities, which are essential in the rest of the proof. We define $\Delta \in \mathbb{R}_{>0}$, the biggest positive constant such that $\{x \in \mathbb{R}^{n_x} : \underline{\alpha}_Y(\sigma(x)) \leq \underline{\alpha}_\varepsilon(\tilde{\varepsilon}) + \Delta\} \subseteq \mathcal{X}$, which exists in view of the definition of $\tilde{\varepsilon}$ above, Assumption 1, and the fact that \mathcal{X} is compact. We also define $\tilde{\delta} := \min \left\{ \left(\mathbb{I} - \frac{\tilde{\alpha}_Y}{2} \right)^{-1} \circ \underline{\alpha}_Y(\delta), \frac{1}{2} \underline{\alpha}_Y\left(\frac{1}{2}\delta\right), \Delta \right\}$ where $\tilde{\alpha}_Y(s) := \alpha_Y \circ \bar{\alpha}_Y^{-1}\left(\frac{1}{2}s\right)$ for any $s \geq 0$. Note that

$(\mathbb{I} - \frac{\underline{\alpha}_Y}{2})^{-1}$ is well-defined as we can assume without loss of generality that $\mathbb{I} - \frac{\underline{\alpha}_Y}{2} \in \mathcal{K}_\infty$, see [8, footnote 5].

We select $\gamma^* \in (0, 1)$ and $\epsilon^* \in (0, \tilde{\epsilon})$ such that

$$\begin{cases} \Upsilon \left(\frac{1-\gamma^*}{\gamma^*} + \epsilon^*, \underline{\alpha}_Y^{-1}(s + \underline{\alpha}_\epsilon(\epsilon^*)) \right) \leq \frac{1}{2} \tilde{\alpha}_Y(s) \\ \forall s \in [\tilde{\delta}, \Delta] \\ \underline{\alpha}_Y^{-1}(2\underline{\alpha}_\epsilon(\epsilon^*)) + \underline{\alpha}_Y^{-1}(2\beta_Y(4\bar{\alpha}_\epsilon(\epsilon^*), 0)) \leq \frac{1}{2} \tilde{\delta} \\ 2\bar{\alpha}_\epsilon(\epsilon^*) + \Upsilon \left(\frac{1-\gamma^*}{\gamma^*} + \epsilon^*, \bar{\alpha}_Y^{-1} \circ \bar{\alpha}_\epsilon(\epsilon^*) \right) \leq \tilde{\delta}, \end{cases} \quad (24)$$

with β_Y some \mathcal{KL} -function defined in the following. This is always possible as Υ is of class \mathcal{KN} , we can thus select a pair $(\gamma^*, \epsilon^*) \in (0, 1) \times (0, \tilde{\epsilon})$ such that $\Upsilon \left(\frac{1-\gamma^*}{\gamma^*} + \epsilon^*, \underline{\alpha}_Y^{-1}(\Delta + \underline{\alpha}_\epsilon(\epsilon^*)) \right) \leq \frac{1}{2} \tilde{\alpha}_Y(\tilde{\delta})$, which ensures the first inequality in (24). The second and third inequalities are satisfied by selecting $\epsilon^* > 0$ sufficiently small and large $\gamma^* \in (0, 1)$ in view of the properties of the involved functions.

Let $\gamma \in [\gamma^*, 1]$, $\epsilon \in [0, \epsilon^*]$, $\varepsilon \in \mathcal{B}_{n_\epsilon}(\epsilon)$, and $x \in \mathcal{X}$ be⁶ such that $Y_{\gamma,\varepsilon}^i(x) \leq \Delta$, and $v \in \hat{F}_{\gamma,\varepsilon,i}(x)$. We first consider the case where $\bar{\alpha}_Y(\sigma(x)) \geq \bar{\alpha}_\varepsilon(|\varepsilon|)$. As a result, in view of item (i) of Theorem 1, $Y_{\gamma,\varepsilon}^i(x) \leq \bar{\alpha}_Y(\sigma(x)) + \bar{\alpha}_\varepsilon(|\varepsilon|) \leq 2\bar{\alpha}_Y(\sigma(x))$. In view of Theorem 1, the definition of $\tilde{\alpha}_Y$ and the fact that $\Upsilon \in \mathcal{KN}$,

$$\begin{aligned} Y_{\gamma,\varepsilon}^i(v) - Y_{\gamma,\varepsilon}^i(x) &\leq -\tilde{\alpha}_Y(Y_{\gamma,\varepsilon}^i(x)) \\ &+ \Upsilon \left(\frac{1-\gamma}{\gamma} + |\varepsilon|, \underline{\alpha}_Y^{-1}(Y_{\gamma,\varepsilon}^i(x) + \underline{\alpha}_\varepsilon(|\varepsilon|)) \right). \end{aligned} \quad (25)$$

Since $\gamma \in [\gamma^*, 1]$, $\varepsilon \in \mathcal{B}_{n_\epsilon}(\epsilon^*)$ and $\Upsilon \in \mathcal{KN}$,

$$\begin{aligned} Y_{\gamma,\varepsilon}^i(v) - Y_{\gamma,\varepsilon}^i(x) &\leq -\tilde{\alpha}_Y(Y_{\gamma,\varepsilon}^i(x)) \\ &+ \Upsilon \left(\frac{1-\gamma^*}{\gamma^*} + \epsilon^*, \underline{\alpha}_Y^{-1}(Y_{\gamma,\varepsilon}^i(x) + \underline{\alpha}_\varepsilon(\epsilon^*)) \right). \end{aligned} \quad (26)$$

When $Y_{\gamma,\varepsilon}^i(x) \in [\tilde{\delta}, \Delta]$, in view of (24) and (26),

$$Y_{\gamma,\varepsilon}^i(v) - Y_{\gamma,\varepsilon}^i(x) \leq -\frac{1}{2} \tilde{\alpha}_Y(Y_{\gamma,\varepsilon}^i(x)). \quad (27)$$

When $Y_{\gamma,\varepsilon}^i(x) \in [0, \tilde{\delta}]$, in view of (24), (26), the definition of $\tilde{\delta}$ and since $\mathbb{I} - \tilde{\alpha}_Y$ can be assumed to be in \mathcal{K}_∞ without loss of generality [8, Remark 5],

$$\begin{aligned} Y_{\gamma,\varepsilon}^i(v) &\leq (\mathbb{I} - \tilde{\alpha}_Y)(Y_{\gamma,\varepsilon}^i(x)) \\ &+ \Upsilon \left(\frac{1-\gamma^*}{\gamma^*} + \epsilon^*, \underline{\alpha}_Y^{-1}(Y_{\gamma,\varepsilon}^i(x) + \underline{\alpha}_\varepsilon(\epsilon^*)) \right) \\ &\leq (\mathbb{I} - \tilde{\alpha}_Y)(\tilde{\delta}) + \Upsilon \left(\frac{1-\gamma^*}{\gamma^*} + \epsilon^*, \underline{\alpha}_Y^{-1}(\tilde{\delta} + \underline{\alpha}_\varepsilon(\epsilon^*)) \right) \\ &\leq (\mathbb{I} - \tilde{\alpha}_Y)(\tilde{\delta}) + \frac{1}{2} \tilde{\alpha}_Y(\tilde{\delta}) \\ &= (\mathbb{I} - \frac{1}{2} \tilde{\alpha}_Y)(\tilde{\delta}). \end{aligned} \quad (28)$$

Thus, whenever $Y_{\gamma,\varepsilon}^i(x) \leq \tilde{\delta}$, $Y_{\gamma,\varepsilon}^i(v) \leq \tilde{\delta}$.

We now consider the case where $\bar{\alpha}_Y(\sigma(x)) \leq \bar{\alpha}_\varepsilon(|\varepsilon|)$, which implies $Y_{\gamma,\varepsilon}^i(x) \leq 2\bar{\alpha}_\varepsilon(|\varepsilon|)$ according to item (i) of Theorem 1. Consequently, in view of item (ii) of Theorem

⁶It is always possible to find such x in view of Assumption 3 and item (i) of Theorem 1.

1,

$$\begin{aligned} Y_{\gamma,\varepsilon}^i(v) &\leq Y_{\gamma,\varepsilon}^i(x) - \alpha_Y(\sigma(x)) \\ &+ \Upsilon \left(\frac{1-\gamma}{\gamma} + |\varepsilon|, \sigma(x) \right) \\ &\leq Y_{\gamma,\varepsilon}^i(x) + \Upsilon \left(\frac{1-\gamma}{\gamma} + |\varepsilon|, \sigma(x) \right) \\ &\leq 2\bar{\alpha}_\varepsilon(|\varepsilon|) \\ &+ \Upsilon \left(\frac{1-\gamma}{\gamma} + |\varepsilon|, \bar{\alpha}_Y^{-1} \circ \bar{\alpha}_\varepsilon(|\varepsilon|) \right). \end{aligned} \quad (29)$$

Since $|\varepsilon| \leq \epsilon^*$ and $\gamma \in [\gamma^*, 1]$,

$$\begin{aligned} Y_{\gamma,\varepsilon}^i(v) &\leq 2\bar{\alpha}_\varepsilon(\epsilon^*) \\ &+ \Upsilon \left(\frac{1-\gamma^*}{\gamma^*} + \epsilon^*, \bar{\alpha}_Y^{-1} \circ \bar{\alpha}_\varepsilon(\epsilon^*) \right) \end{aligned} \quad (30)$$

We then deduce from the last inequality in (24) that

$$\max\{Y_{\gamma,\varepsilon}^i(x), Y_{\gamma,\varepsilon}^i(v)\} \leq \tilde{\delta}. \quad (31)$$

The solutions to (6) initialized at x remain in the set \mathcal{X} for all positive times. Indeed, since $x \in \mathcal{X}$ and $Y_{\gamma,\varepsilon}^i(x) \leq \Delta$, $Y_{\gamma,\varepsilon}^i(v) \leq Y_{\gamma,\varepsilon}^i(x) \leq \Delta$ according to (27). Moreover, (28) and (31) imply that $Y_{\gamma,\varepsilon}^i(v) \leq \tilde{\delta} \leq \Delta$ by definition of $\tilde{\delta}$. Hence, in both cases, $Y_{\gamma,\varepsilon}^i(v) \leq \Delta$ which means that $v \in \mathcal{X}$ as $\{z \in \mathbb{R}^{n_x} : Y_{\gamma,\varepsilon}^i(z) \leq \Delta\} \subset \{z \in \mathbb{R}^{n_x} : \underline{\alpha}_Y(\sigma(z)) \leq \underline{\alpha}_\varepsilon(\tilde{\epsilon}) + \Delta\} \subset \mathcal{X}$ in view of item (i) of Theorem 1 and the definition of Δ . By proceeding iteratively, we deduce that any solution to (6) initialized at x remain in the set \mathcal{X} .

Based on the above fact, item (i) of Theorem 1, (27) and (28), we follow the same arguments as in the proof of Theorem 2 in [21] to conclude that there exists $\tilde{\beta}_Y \in \mathcal{KL}$ such that for any $k \in \mathbb{Z}_{\geq 0}$ and any solution ϕ to system (6) initialized at x ,

$$Y_{\gamma,\varepsilon}^i(\phi(k, x)) \leq \max\{\beta_Y(Y_{\gamma,\varepsilon}^i(x), k), \tilde{\delta}\}. \quad (32)$$

We deduce from item (i) of Theorem 1, since $|\varepsilon| \leq \epsilon^*$, $\underline{\alpha}_Y(\sigma(\phi(k, x))) - \underline{\alpha}_\varepsilon(\epsilon^*) \leq \max\{\beta_Y(\bar{\alpha}_Y(\sigma(x)) + \bar{\alpha}_\varepsilon(\epsilon^*), k), \tilde{\delta}\}$. Hence, $\sigma(\phi(k, x)) \leq \underline{\alpha}_Y^{-1} \left(\max\{\beta_Y(\bar{\alpha}_Y(\sigma(x)) + \bar{\alpha}_\varepsilon(\epsilon^*), k), \tilde{\delta}\} + \underline{\alpha}_\varepsilon(\epsilon^*) \right)$. Using the property that $\alpha(s_1 + s_2) \leq \alpha(2s_1) + \alpha(2s_2)$ for any $\alpha \in \mathcal{K}$, $s_1, s_2 \in \mathbb{R}_{\geq 0}$, see [13, (6)], we derive $\sigma(\phi(k, x)) \leq \underline{\alpha}_Y^{-1} \left(\max\{\beta_Y(2\bar{\alpha}_Y(\sigma(x)), k) + \beta_Y(2\bar{\alpha}_\varepsilon(\epsilon^*), 0), \tilde{\delta}\} + \underline{\alpha}_\varepsilon(\epsilon^*) \right)$. Thus, by using twice the same property $\sigma(\phi(k, x)) \leq \underline{\alpha}_Y^{-1} \left(2 \max\{\beta_Y(2\bar{\alpha}_Y(\sigma(x)), k) + \beta_Y(2\bar{\alpha}_\varepsilon(\epsilon^*), 0), \tilde{\delta}\} \right) + \underline{\alpha}_Y^{-1}(2\underline{\alpha}_\varepsilon(\epsilon^*)) = \max\{\underline{\alpha}_Y^{-1}(2\beta_Y(2\bar{\alpha}_Y(\sigma(x)), k) + 2\beta_Y(2\bar{\alpha}_\varepsilon(\epsilon^*), 0)), \underline{\alpha}_Y^{-1}(2\tilde{\delta})\} + \underline{\alpha}_Y^{-1}(2\underline{\alpha}_\varepsilon(\epsilon^*)) \leq \max\{\underline{\alpha}_Y^{-1}(4\beta_Y(2\bar{\alpha}_Y(\sigma(x)), k)) + \underline{\alpha}_Y^{-1}(4\beta_Y(2\bar{\alpha}_\varepsilon(\epsilon^*), 0)), \underline{\alpha}_Y^{-1}(2\tilde{\delta})\} + \underline{\alpha}_Y^{-1}(2\underline{\alpha}_\varepsilon(\epsilon^*))$. Since $\max\{s_1, s_2\} \leq s_1 + s_2$ for any $s_1, s_2 \in \mathbb{R}_{\geq 0}$,

$$\begin{aligned} \sigma(\phi(k, x)) &\leq \underline{\alpha}_Y^{-1}(4\beta_Y(2\bar{\alpha}_Y(\sigma(x)), k)) \\ &+ \underline{\alpha}_Y^{-1}(4\beta_Y(2\bar{\alpha}_\varepsilon(\epsilon^*), 0)) + \underline{\alpha}_Y^{-1}(2\tilde{\delta}) \\ &+ \underline{\alpha}_Y^{-1}(2\underline{\alpha}_\varepsilon(\epsilon^*)). \end{aligned} \quad (33)$$

By definition of $\tilde{\delta}$ and in view of the second inequality in (24),

$$\sigma(\phi(k, x)) \leq \underline{\alpha}_Y^{-1}(4\beta_Y(2\bar{\alpha}_Y(\sigma(x)), k)) + \delta. \quad (34)$$

This inequality ensures that (12) is satisfied with $\beta(s_1, s_2) = \underline{\alpha}_Y^{-1}(4\beta_Y(2\bar{\alpha}_Y(s_1), s_2))$ for any $s_1, s_2 \geq 0$. ■

Sketch of proof of Corollary 1. As in the proof of Theorem 2, let $\tilde{\epsilon} \in (0, \bar{\epsilon}]$ be small enough such that the set $\{x \in \mathbb{R}^{n_x} : \underline{\alpha}_Y(\sigma(x)) \leq \underline{\alpha}_\epsilon(\tilde{\epsilon})\}$ is included in the interior of \mathcal{X} . We define $\Delta \in \mathbb{R}_{>0}$, the biggest positive constant such that $\{x \in \mathbb{R}^{n_x} : \underline{\alpha}_Y(\sigma(x)) \leq \underline{\alpha}_\epsilon(\tilde{\epsilon}) + \Delta\} \subseteq \mathcal{X}$, which exists in view of the definition of $\tilde{\epsilon}$ above, Assumption 1, and the fact that \mathcal{X} is compact.

Let $i \geq i^*$, $\gamma \in [\gamma^*, 1]$, $\epsilon \in [0, \epsilon^*]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\epsilon)$, $x \in \mathcal{X}$ such that $Y_{\gamma, \varepsilon}^i(x) \leq \Delta$ and $v \in \hat{F}_{\gamma, \varepsilon, i}(x)$. In view of (21), (23) and items (ii)-(iii) of Corollary 1,

$$\begin{aligned} Y_{\gamma, \varepsilon}^i(v) - Y_{\gamma, \varepsilon}^i(x) &\leq -a_W \sigma(x) \\ &\quad + \frac{1-\gamma}{\gamma} (\bar{a}_V \sigma(x) + a_{\text{stop}}(|\varepsilon|) \sigma(x)) \\ &\quad + a_{\text{stop}}(|\varepsilon|) \sigma(x) \\ &= \left(-a_W + \frac{1-\gamma}{\gamma} (\bar{a}_V + a_{\text{stop}}(2|\varepsilon|)) \right. \\ &\quad \left. + a_{\text{stop}}(2|\varepsilon|) \right) \sigma(x). \end{aligned} \quad (35)$$

Since $\gamma \in [\gamma^*, 1]$ and $\epsilon \in [0, \epsilon^*]$, $-a_W + \frac{1-\gamma}{\gamma} (\bar{a}_V + a_{\text{stop}}(|\varepsilon|)) + a_{\text{stop}}(|\varepsilon|) \leq -a_W + \frac{1-\gamma^*}{\gamma^*} (\bar{a}_V + a_{\text{stop}}(\epsilon^*)) + a_{\text{stop}}(\epsilon^*) < 0$. Hence, there exists $\nu > 0$ independent of (γ, ϵ) and sufficiently small such that

$$Y_{\gamma, \varepsilon}^i(v) - Y_{\gamma, \varepsilon}^i(x) \leq -\nu \sigma(x). \quad (36)$$

We then apply similar arguments as in the proof of Theorem 2 to obtain the desired result.

When, in addition $\underline{\alpha}_W(s) = \underline{\alpha}_W \cdot s$ for any $s \geq 0$, we also have that $\underline{\alpha}_Y \sigma(x) \leq Y_{\gamma, \varepsilon}^i(x) \leq \bar{a}_Y \sigma(x)$ in view of item (i) of Theorem 1 and item (iii) of Corollary 1. We then deduce from (36) that $\beta \in \exp -\mathcal{KL}$ in (14), like in the proof of Corollary 2 in [21]. ■

Proof of Corollary 2. We first prove that the standing assumptions hold. Let $x \in \mathcal{X}$, u be any admissible input, $\varepsilon_V \in \mathcal{B}_{n_V}(\bar{\varepsilon})$ and $i \in \mathbb{Z}_{\geq 0}$. In view of item (iv) of Corollary 2, $\ell(x, u) + e_{\hat{V}, i+1}(1, \varepsilon_V, x) \geq \ell(x, u) - |\varepsilon_V| \ell(x, 0) \geq (1-\bar{\varepsilon})Q(x) + R(u)$. Recalling that Q and R take non-negative values and $\bar{\varepsilon} \in [0, 1)$ according to item (iv) of Corollary 2, we derive that $\hat{V}_{1, \varepsilon_V}^i(x) \geq 0$. SA1 follows by invoking item (v) of Corollary 2.

SA2, and item (ii) of Corollary 1, trivially hold in view of item (vi) of Corollary 2.

In view of items (iv) and (vi) of Corollary 2, we can invoke [12, (20), (37)], that is, for any $x \in \mathcal{X}$, $\varepsilon_V \in \mathcal{B}_{n_V}(\bar{\varepsilon})$ and $i \in \mathbb{Z}_{\geq 0}$,

$$\begin{aligned} \mathcal{T}\hat{V}_{1, \varepsilon_V}^i(x) &\leq \hat{V}_{1, \varepsilon_V}^i(x) + 2|\varepsilon_V| \check{V}^i(x) \\ &\leq \hat{V}_{1, \varepsilon_V}^i(x) + \frac{2|\varepsilon_V|}{1-|\varepsilon_V|} \bar{V}^0(x), \end{aligned} \quad (37)$$

where $\check{V}^i(x)$ is defined in [12, Lemma 1] and $\bar{V}^0(x) \leq 2\hat{V}^0(x)$. Therefore, in view of item (vii) of Corollary 2,

$$\mathcal{T}\hat{V}_{1, \varepsilon_V}^i(x) \leq \hat{V}_{1, \varepsilon_V}^i(x) + \frac{2\bar{\varepsilon}}{1-\bar{\varepsilon}} 2dQ(x). \quad (38)$$

Thus SA3 holds for $x \in \mathcal{X}$ with $e_{\text{stop}, i}(1, \varepsilon_V, x) = \frac{4\varepsilon_V}{1-\varepsilon_V} dQ(x)$ for any $x \in \mathcal{X}$, any $i \in \mathbb{Z}_{\geq 0}$ and there is no parameter $\varepsilon_{\text{stop}}$. Also, SA3 holds is trivially verified with $e_{\text{stop}, i}(1, \varepsilon_V, x) = \mathcal{T}\hat{V}_{\gamma, \varepsilon_V}^i(x)$ when $x \in \mathbb{R}^{n_x} \setminus \mathcal{X}$. Note that $i^* = 0$ here and that ε reduces to ε_V .

We prove in the following that Assumptions 1-5 are satisfied.

Assumption 1 is verified by taking $\sigma = Q$, as Q is continuous according to item (ii) of Corollary 2.

In view of item (ii) of Corollary 2, Assumption 2 is verified with $W = 0$, $\bar{\alpha}_W = \alpha_W = a_W \mathbb{I}$ with $a_W = 1$ as $\sigma = Q$ and $\ell(x, u) \geq \ell(x, 0) = Q(x) = \sigma(x)$ for any $x \in \mathbb{R}^{n_x}$ and $u \in \mathbb{R}^{n_u}$.

Assumption 3 is satisfied in view of items (i)-(ii) of Corollary 2 and the definition of σ . Hence, item (i) of Corollary 1 holds.

The first inequality in (9) is verified in view of (38) by taking $\alpha_{\text{stop}}(|\varepsilon|, \sigma(x)) = \frac{4\varepsilon}{1-\varepsilon} d\sigma(x)$ for any $x \in \mathcal{X}$ and $\varepsilon \in \mathcal{B}_{n_V}(\bar{\varepsilon})$, which indeed defines a \mathcal{KN} -function. Moreover, $\alpha_{\text{stop}}(|\varepsilon|, \sigma(x)) = a_{\text{stop}}(|\varepsilon|) \sigma(x)$ with $a_{\text{stop}}(s) = \frac{4s}{1-s} d$ for any $s \in [0, 1)$, which defines a class- \mathcal{K} function as required in item (iii) of Corollary 1. The third and fourth inequalities in Assumption 4 and item (ii) of Corollary 1 hold as $e_{\hat{h}, \varepsilon_{h,i}} = 0$ in view of item (vi) of Corollary 2. Regarding (9b), for any $x \in \mathcal{X}$, $i \in \mathbb{Z}_{\geq 0}$ and $\varepsilon \in \mathcal{B}_{n_V}(\bar{\varepsilon})$, we have $|e_{\hat{V}, i}(1, \varepsilon, x)| \leq |\varepsilon| \ell(x, 0) = Q(x) - (1-|\varepsilon|)Q(x) \leq Q(x) - (1-\bar{\varepsilon})Q(x)$. Therefore, (9b) holds with $\underline{\alpha}_W = \underline{\alpha}_W \mathbb{I}$, $\underline{a}_W = 1 - \bar{\varepsilon}$ and $\underline{\alpha}_\varepsilon = 0$.

Let $x \in \mathcal{X}$, $\varepsilon \in \mathcal{B}_{n_V}(\bar{\varepsilon})$ and $i \geq 0$. As $\hat{V}_{1, \varepsilon}^i(x) \leq \bar{V}^i(x)$ according to [12, Theorem 1] and $\bar{V}^i(x) \leq \bar{V}^0(x)$ as explained after [12, (36)], $\hat{V}_{1, \varepsilon}^i(x) \leq \bar{V}^0(x)$ and since $\bar{V}^0(x) \leq 2dQ(x)$ as shown above, $\hat{V}_{1, \varepsilon}^i(x) \leq 2dQ(x)$. Thus, (10) is verified with $\bar{\alpha}_V = 2d\mathbb{I} \in \mathcal{K}_{\mathcal{X}}$ and $\bar{\alpha}_\varepsilon = 0$: Assumption 5 is satisfied.

We have proved that item (iii) of Corollary 1 is satisfied. This completes the proof. ■

REFERENCES

- [1] D. P. Bertsekas and J. N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, Belmont, U.S.A., 1996.
- [2] D.P. Bertsekas. Dynamic programming and suboptimal control: A survey from ADP to MPC. *European Journal of Control*, 11(4-5):310–334, 2005.
- [3] D.P. Bertsekas. *Abstract Dynamic Programming*. Athena S, 2nd edition, 2018.
- [4] T. Bian and Z.-P. Jiang. Value iteration and adaptive dynamic programming for data-driven adaptive optimal control design. *Automatica*, 71:348–360, 2016.
- [5] L. Buşoniu, D. Ernst, B. De Schutter, and R. Babuška. Approximate dynamic programming with a fuzzy parameterization. *Automatica*, 46(5):804–814, 2010.
- [6] A. M. Farahmand, M. Ghavamzadeh, C. Szepesvári, and S. Mannor. Regularized fitted Q-iteration for planning in continuous-space Markovian decision problems. In *American Control Conference, St. Louis, U.S.A.*, pages 725–730, 2009.

-
- [7] M. Granzotto, R. Postoyan, L. Buşoniu, D. Nešić, and J. Daafouz. Stability analysis of discrete-time finite-horizon discounted optimal control. In *IEEE Conf. on Dec. and Control, Miami, U.S.A.*, 2018.
- [8] M. Granzotto, R. Postoyan, L. Buşoniu, D. Nešić, and J. Daafouz. Finite-horizon discounted optimal control: stability and performance. *submitted for journal publication*, 2019.
- [9] G. Grimm, M.J. Messina, S.E. Tuna, and A.R. Teel. Model predictive control: for want of a local control Lyapunov function, all is not lost. *IEEE Transactions on Automatic Control*, 50(5):546–558, 2005.
- [10] A. Heydari. Theoretical and numerical analysis of approximate dynamic programming with approximation errors. *Journal of Guidance, Control, and Dynamics*, 39(2):301–311, 2015.
- [11] A. Heydari. Stability analysis of optimal adaptive control under value iteration using a stabilizing initial policy. *IEEE Transactions on Neural Networks and Learning Systems*, 29(9):4522–4527, 2018.
- [12] A. Heydari. Stability analysis of optimal adaptive control using value iteration with approximation errors. *IEEE Transactions on Automatic Control*, 63(9):3119 – 3126, 2018.
- [13] Z.-P. Jiang, A.R. Teel, and L. Praly. Small-gain theorem for ISS systems and applications. *Mathematics of Control, Signals, and Systems*, 7:95–120, 1994.
- [14] C.M. Kellett and A.R. Teel. On the robustness of \mathcal{KL} -stability for difference inclusions: smooth discrete-time Lyapunov functions. *SIAM Journal on Control and Optimization*, 44(3):777–800, 2005.
- [15] H.K. Khalil. *Nonlinear systems*. Prentice-Hall, Englewood Cliffs, New Jersey, U.S.A., 3rd edition, 2002.
- [16] D. Liu and Q. Wei. Finite-approximation-error-based optimal control approach for discrete-time nonlinear systems. *IEEE Transactions on Cybernetics*, 43(2):779–789, 2013.
- [17] D. Liu, Q. Wei, D. Wang, X. Yang, and H. Li. Finite approximation error-based value iteration ADP. In *Adaptive Dynamic Programming with Applications in Optimal Control*, pages 91–149. Springer, 2017.
- [18] D. Liu, Q. Wei, D. Wang, X. Yang, and H. Li. Value iteration ADP for discrete-time nonlinear systems. In *Adaptive Dynamic Programming with Applications in Optimal Control*, pages 37–90. Springer, 2017.
- [19] R. Munos and C. Szepesvári. Finite-time bounds for fitted value iteration. *Journal of Machine Learning Research*, 9:815–857, 2008.
- [20] R. Postoyan. *Commande et construction d’observateurs pour les systèmes non linéaires à données échantillonnées et en réseau*. PhD thesis, Univ Paris-Sud (in French), 2009.
- [21] R. Postoyan, L. Buşoniu, D. Nešić, and J. Daafouz. Stability analysis of discrete-time infinite-horizon optimal control with discounted cost. *IEEE Transactions on Automatic Control*, 62(6):2736–2749, 2017.
- [22] M. Rinehart, M. Dahleh, and I. Kolmanovsky. Value iteration for (switched) homogeneous systems. *IEEE Transactions on Automatic Control*, 54(6):1290–1294, 2009.
- [23] B. Scherrer and B. Lesner. On the use of non-stationary policies for stationary infinite-horizon Markov decision processes. In *NIPS - Neur. Inf. Processing Syst.*, South Lake Tahoe, U.S.A., 2012.
- [24] S.P. Singh and R.C. Yee. An upper bound on the loss from approximate optimal-value functions. *Machine Learning*, 16(3):227–233, 1994.
- [25] A.R. Teel and L. Praly. A smooth Lyapunov function from a class- \mathcal{KL} estimate involving two positive semidefinite functions. *ESAIM: Control, Optimisation and Calculus of Variations*, 5(1):313–367, 2000.
- [26] Q. Wei and D. Liu. Stable iterative adaptive dynamic programming algorithm with approximation errors for discrete-time nonlinear systems. *Neural Computing and Applications*, 24(6):1355–1367, 2014.
- [27] Q. Wei, D. Liu, and H. Lin. Value iteration adaptive dynamic programming for optimal control of discrete-time nonlinear systems. *IEEE Transactions on Cybernetics*, 46(3):840–853, 2016.
- [28] R.J. Williams and L.C. Baird. Tight performance bounds on greedy policies based on imperfect value functions. In *Proceedings of the 8th Yale Workshop on Adaptive and Learning Systems, New Haven, U.S.A.*, pages 108–113, 1994.

Bibliography

- [1] A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf. Discrete-time nonlinear hjb solution using approximate dynamic programming: Convergence proof. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 38(4):943–949, 2008.
- [2] B. D. O. Anderson and J. B. Moore. *Optimal control: linear quadratic methods*. Courier Corporation, 2007.
- [3] D. Antunes and W. P. M. H. Heemels. Linear quadratic regulation of switched systems using informed policies. *IEEE Transactions on Automatic Control*, 62(6):2675–2688, 2017.
- [4] G. I. Bara and M. Boutayeb. A new sufficient condition for the static output feedback stabilization of linear discrete-time systems. In *IEEE Conference on Decision and Control*, pages 4723–4728, San Diego, USA, 2006.
- [5] R. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, NJ, USA, 2010.
- [6] D. P. Bertsekas. Dynamic programming and suboptimal control: A survey from adp to mpc. *European Journal of Control*, 11(4-5):310–334, 2005.
- [7] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 2. Athena Scientific, Nashua, USA, 4th edition, 2012.
- [8] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, Nashua, USA, 4th edition, 2012.

- [9] D. P. Bertsekas. Value and policy iterations in optimal control and adaptive dynamic programming. *IEEE Transactions on Neural Networks and Learning Systems*, 28(3):500–509, 2017.
- [10] D. P Bertsekas and J. N. Tsitsiklis. *Neuro-dynamic programming*, volume 5. Athena Scientific Belmont, MA, 1996.
- [11] S. Bhasin, R. Kamalapurkar, M. Johnson, K.G. Vamvoudakis, F.L. Lewis, and W.E. Dixon. A novel actor–critic–identifier architecture for approximate optimal control of uncertain nonlinear systems. *Automatica*, 49(1):82 – 92, 2013.
- [12] T. Bian, Y. Jiang, and Z.-P. Jiang. Adaptive dynamic programming and optimal control of nonlinear nonaffine systems. *Automatica*, 50(10):2624–2632, 2014.
- [13] C. I. Boussios, M. A. Dahleh, and J. N. Tsitsiklis. Semiglobal nonlinear stabilization via approximate policy iteration. In *American Control Conference*, volume 6, pages 4675–4680 vol.6, Arlington, VA, USA, 2001.
- [14] L. Buşoniu, R. Munos, B. De Schutter, and R. Babuška. Optimistic planning for sparsely stochastic systems. In *IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)*, pages 48–55, Paris, France, 2011.
- [15] L. Buşoniu, J. Daafouz, M. C. Bragagnolo, and I.-C. Morărescu. Planning for optimal control and performance certification in nonlinear systems with controlled or uncontrolled switches. *Automatica*, 78:297 – 308, 2017.
- [16] L. Buşoniu, D. Ernst, B. D. Schutter, and R. Babuška. Approximate dynamic programming with a fuzzy parameterization. *Automatica*, 46(5):804 – 814, 2010.
- [17] L. Buşoniu, R. Postoyan, and J. Daafouz. Near-optimal strategies for nonlinear and uncertain networked control systems. *IEEE Transactions on Automatic Control*, 61(8):2124–2139, 2016.
- [18] X. Chang and G. Yang. New results on output feedback h_∞ control for linear discrete-time systems. *IEEE Transactions on Automatic Control*, 59(5):1355–1359, 2014.

-
- [19] G. Feng, E. Meyer, and Y. Liu. A new digital control algorithm to achieve optimal dynamic performance in dc-to-dc converters. *IEEE Transactions on Power Electronics*, 22(4):1489–1498, 2007.
- [20] C. A. C. Gonzaga, M. Jungers, and J. Daafouz. Stability analysis of discrete-time Lur’e systems. *Automatica*, 48(9):2277 – 2283, 2012.
- [21] M. Granzotto, R. Postoyan, L. Buşoniu, D. Nešić, and J. Daafouz. Finite-horizon discounted optimal control: stability and performance. In *submitted for journal publication*.
- [22] M. Granzotto, R. Postoyan, L. Buşoniu, D. Nešić, and J. Daafouz. Stability analysis of discrete-time finite-horizon optimal control with discounted cost. In *IEEE Conference on Decision and Control*, pages 2322–2327, Miami, USA, 2018.
- [23] M. Granzotto, R. Postoyan, L. Buşoniu, D. Nešić, and J. Daafouz. Optimistic planning for the near-optimal control of nonlinear switched discrete-time systems with stability guarantees. In *IEEE Conference on Decision and Control*, Nice, France, 2019. An arxiv version can be found at: <https://arxiv.org/pdf/1908.01404.pdf>.
- [24] G. Grimm, M. J. Messina, S. E. Tuna, and A. R. Teel. Model predictive control: for want of a local control Lyapunov function, all is not lost. *IEEE Transactions on Automatic Control*, 50(5):546–558, 2005.
- [25] L. Grüne and A. Rantzer. On the infinite horizon performance of receding horizon controllers. *IEEE Transactions on Automatic Control*, 53(9):2100–2111, 2008.
- [26] L. Grüne, E. D. Sontag, and F. R. Wirth. Asymptotic stability equals exponential stability, and iss equals finite energy gain — if you twist your eyes. *Systems & Control Letters*, 38(2):127 – 134, 1999.
- [27] A. Heydari. Revisiting approximate dynamic programming and its convergence. *IEEE Transactions on Cybernetics*, 44(12):2733–2743, 2014.

- [28] A. Heydari. Theoretical and numerical analysis of approximate dynamic programming with approximation errors. *Journal of Guidance, Control, and Dynamics*, 39(2):301–311, 2015.
- [29] A. Heydari. Analyzing policy iteration in optimal control. In *IEEE American Control Conference*, pages 5728–5733, 2016.
- [30] A. Heydari. Stability analysis of optimal adaptive control under value iteration using a stabilizing initial policy. *IEEE Transactions on Neural Networks and Learning Systems*, 29(9):4522–4527, 2017.
- [31] A. Heydari. Stability analysis of optimal adaptive control using value iteration with approximation errors. *IEEE Transactions on Automatic Control*, 63(9):3119–3126, 2018.
- [32] J.-F. Hren and R. Munos. Optimistic planning of deterministic systems. In *European Workshop on Reinforcement Learning*, pages 151–164, Villeneuve d’Ascq, France, 2008.
- [33] M. Höger and L. Grüne. On the relation between detectability and strict dissipativity for nonlinear discrete time systems. *IEEE Control Systems Letters*, 3(2):458–462, 2019.
- [34] S. Ibrir. Static output feedback and guaranteed cost control of a class of discrete-time nonlinear systems with partial state measurements. *Nonlinear Analysis: Theory, Methods & Applications*, 68(7):1784 – 1792, 2008.
- [35] Y. Jiang and Z.-P. Jiang. Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics. *Automatica*, 48(10):2699–2704, 2012.
- [36] Y. Jiang and Z.-P. Jiang. Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics. *Automatica*, 48(10):2699 – 2704, 2012.
- [37] Z.-P. Jiang and Y. Wang. Input-to-state stability for discrete-time nonlinear systems. *Automatica*, 37(6):857–869, 2001.

-
- [38] R.E. Kalman. Contributions to the theory of optimal control. *Bol. soc. mat. mexicana*, 5(2):102–119, 1960.
- [39] S. Keerthi and E. Gilbert. An existence theorem for discrete-time infinite-horizon optimal control problems. *IEEE Transactions on Automatic Control*, 30(9):907–909, 1985.
- [40] C. M. Kellett and A. R. Teel. On the robustness of \mathcal{KL} -stability for difference inclusions: Smooth discrete-time Lyapunov functions. *SIAM Journal on Control and Optimization*, 44(3):777–800, 2005.
- [41] J. B. Lasserre, D. Henrion, C. Prieur, and E. Trélat. Nonlinear optimal control via occupation measures and lmi-relaxations. *SIAM Journal on Control and Optimization*, 47(4):1643–1666, 2008.
- [42] F.L. Lewis and D. Vrabie. Reinforcement learning and adaptive dynamic programming for feedback control. *Circuits and Systems Magazine, IEEE*, 9:32 – 50, 01 2009.
- [43] D. Liberzon. *Switching in Systems and Control*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Boston, MA, 1st ed. 2003. edition, 2003.
- [44] B. Lincoln and A. Rantzer. Relaxing dynamic programming. *IEEE Transactions on Automatic Control*, 51(8):1249–1260, Aug 2006.
- [45] D. Liu and Q. Wei. Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems. *IEEE Transactions on Neural Networks and Learning Systems*, 25(3):621–634, 2013.
- [46] D. Liu, Q. Wei, D. Wang, X. Yang, and H. Li. Finite approximation error-based value iteration ADP. In *Adaptive Dynamic Programming with Applications in Optimal Control*, pages 91–149. Springer, 2017.
- [47] D. Liu, Q. Wei, D. Wang, X. Yang, and H. Li. Value iteration ADP for discrete-time nonlinear systems. In *Adaptive Dynamic Programming with Applications in Optimal Control*, pages 37–90. Springer, 2017.

- [48] B. Luo, D. Liu, T. Huang, and D. Wang. Model-free optimal tracking control via critic-only Q-learning. *IEEE Transactions on neural networks and learning systems*, 27(10):2134–2144, 2016.
- [49] D. Q. Mayne. Model predictive control: Recent developments and future promise. *Automatica*, 50(12):2967–2986, 2014.
- [50] J. D. Meiss. *Differential Dynamical Systems (Monographs on Mathematical Modeling and Computation)*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2007.
- [51] F. Morbidi, R. Cano, and D. Lara. Minimum-energy path generation for a quadrotor uav. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 1492–1498, 2016.
- [52] R. Munos and C. Szepesvári. Finite time bounds for fitted value iteration. *Journal of Machine Learning Research*, 9:815–857, 2008.
- [53] J. J. Murray, C. J. Cox, G. G. Lendaris, and R. Saeks. Adaptive dynamic programming. *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, 32(2):140–153, 2002.
- [54] M. A. Müller and L. Grüne. On the relation between dissipativity and discounted dissipativity. In *IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 5570–5575, Melbourne, AU, 2017.
- [55] R. Postoyan, L. Buşoniu, D. Nešić, and J. Daafouz. Stability analysis of discrete-time infinite-horizon optimal control with discounted cost. *IEEE Transactions on Automatic Control*, 62(6):2736–2749, 2017.
- [56] R. Postoyan, M. Granzotto, L. Buşoniu, B. Scherrer, D. Nešić, and J. Daafouz. Stability guarantees for nonlinear discrete-time systems controlled by approximate value iteration. In *IEEE Conference on Decision and Control*, Nice, France, 2019.
- [57] C. Prieur. Uniting local and global controllers with robustness to vanishing noise. *Mathematics of Control, Signals and Systems*, 14(2):143–172, 2001.

-
- [58] C. Prieur and L. Praly. Uniting local and global controllers. In *IEEE Conference on Decision and Control*, volume 2, pages 1214–1219, Phoenix, USA, 1999.
- [59] C. Prieur and A. R. Teel. Uniting local and global output feedback controllers. *IEEE Transactions on Automatic Control*, 56(7):1636–1649, 2011.
- [60] J. B. Rejeb, L. Buşoniu, I. Morărescu, and J. Daafouz. Near-optimal control of nonlinear switched systems with non-cooperative switching rules. In *American Control Conference*, pages 2648–2653, Seattle, WA, USA, 2017.
- [61] A. Richards, T. Schouwenaars, J. P. How, and E. Feron. Spacecraft trajectory planning with avoidance constraints using mixed-integer linear programming. *Journal of Guidance, Control, and Dynamics*, 25(4):755–764, 2002.
- [62] P. Riedinger. A switched LQ regulator design in continuous time. *IEEE Transactions on Automatic Control*, 59(5):1322–1328, 2014.
- [63] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 2009.
- [64] C. Savorgnan, J. B. Lasserre, and M. Diehl. Discrete-time stochastic optimal control via occupation measures and moment relaxations. In *Proceedings of the 48th IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 519–524. IEEE, 2009.
- [65] B. Scherrer and B. Lesner. On the use of non-stationary policies for stationary infinite-horizon markov decision processes. In *Advances in Neural Information Processing Systems*, pages 1826–1834, 2012.
- [66] S. P. Singh and R. C. Yee. An upper bound on the loss from approximate optimal-value functions. *Machine Learning*, 16(3):227–233, 1994.
- [67] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, USA, 2nd edition, 2017.

- [68] E. Trélat. *Contrôle optimal : théorie & applications*. Vuibert, 2005. ISBN 2 7117 7175 X.
- [69] S. E. Tuna, M. J. Messina, and A. R. Teel. Shorter horizons for model predictive control. In *American Control Conference*, pages 863–868, Minneapolis, USA, 2006.
- [70] K. G. Vamvoudakis and F. L. Lewis. Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica*, 46(5):878 – 888, 2010.
- [71] V. I. Vorotnikov. *Partial stability and control*. Springer Science & Business Media, 2012.
- [72] D. Vrabie and F. Lewis. Neural network approach to continuous-time direct adaptive optimal control for partially unknown nonlinear systems. *Neural Networks*, 22(3):237 – 246, 2009. Goal-Directed Neural Systems.
- [73] D. Vrabie, O. Pastravanu, M. Abu-Khalaf, and F.L. Lewis. Adaptive optimal control for continuous-time linear systems based on policy iteration. *Automatica*, 45(2):477 – 484, 2009.
- [74] Q. Wei and D. Liu. Stable iterative adaptive dynamic programming algorithm with approximation errors for discrete-time nonlinear systems. *Neural Computing and Applications*, 24(6):1355–1367, 2014.
- [75] Qinglai Wei, Derong Liu, and Hanquan Lin. Value iteration adaptive dynamic programming for optimal control of discrete-time nonlinear systems. *IEEE Transactions on Cybernetics*, 46(3):840–853, 2015.
- [76] M. Wiering and M. van Otterlo, editors. *Reinforcement Learning: State of the Art*, volume 12. Springer, 2012.
- [77] F. Zhang, H. L. Trentelman, G. Feng, and J. M.A. Scherpen. Absolute stabilization of Lur’e systems via dynamic output feedback. *European Journal of Control*, 44:15 – 26, 2018. Advanced Control and Observer Design for Nonlinear Systems via LMIs.

-
- [78] W. Zhang, J. Hu, and A. Abate. A study of the discrete-time switched LQR problem. *ECE TechnicalReports*, (Paper 384), 2009.
- [79] W. Zhang, J. Hu, and A. Abate. Infinite-horizon switched LQR problems in discrete time: A suboptimal algorithm with performance analysis. *IEEE Transactions on Automatic Control*, 57(7):1815–1821, 2012.

Résumé

Le contrôle optimal consiste à concevoir les entrées d'un système dynamique de manière à ce qu'elles optimisent une fonction de coût donnée. Ses applications sont très variées, allant de la planification de la trajectoire des engins spatiaux à l'augmentation de l'endurance des drones en vol, ou encore aux convertisseurs de puissance à découpage plus performants pour ne citer que quelques exemples. La mise en place d'un contrôle optimal remonte au problème de la courbe de Brachistochrone. Proposé en 1696 par J. Bernoulli, le défi consistait à trouver "la courbe tracée par un point agissant uniquement par gravité, qui part de A et atteint B dans le temps le plus court". Résolu par plusieurs mathématiciens de l'époque, ce problème a finalement conduit à la naissance du calcul des variations, et enfin, au principe du maximum de Pontryagin. Il est admis que le contrôle optimal moderne a prospéré après la formulation du principe du maximum de Pontryagin dans les années 1950, et de nombreux outils et applications ont suivi.

Le problème du contrôle optimal est notoirement difficile en général. Seuls quelques cas importants, comme lorsque le système est linéaire et que le coût est quadratique, existent des solutions simples et élégantes. Il est très difficile, voire impossible, de construire explicitement une séquence optimale d'entrées en général lorsque le système est non linéaire et le coût de la scène n'est pas quadratique. C'est le cas de nombreux problèmes, comme ceux de la littérature d'apprentissage par renforcement. Une approche puissante pour surmonter ces difficultés, c'est de recourir à la programmation dynamique, voir l'intelligence artificielle.

L'intelligence artificielle est riche en algorithmes de commande optimale. Ces algorithmes s'appuient sur le principe d'optimalité de Bellman, qui est au centre de la

programmation dynamique. Comme décrit par R.E. Bellman, "Une politique optimale a la propriété que quels que soient l'état initial et la décision initiale, les décisions restantes doivent constituer une politique optimale au regard de l'état résultant de la première décision". Au lieu d'optimiser une séquence infinie d'entrées, le principe de Bellman réduit le problème du contrôle optimal à un problème de décision d'une seule entrée. En contrepartie, la fonction de coût optimale doit être connue. C'est donc ici où plusieurs algorithmes ont été développés pour approcher la fonction de coût optimale. Par le biais d'un facteur de "décompte" ou "oublie", il est possible d'obtenir des bornes d'erreur entre le coût optimal, le désiré, est celui obtenu par le calcul numérique. Ces méthodes sont applicables à de larges classes de systèmes non-linéaires en temps discret et ont fait leurs preuves dans de nombreuses applications. Leur exploitation en automatique s'avère donc très prometteuse. Une question fondamentale reste néanmoins à élucider pour cela: celle de la stabilité. En effet, ces travaux se concentrent sur l'optimalité et ignorent dans la plupart des cas la stabilité du système commandé, qui est au cœur de l'automatique. La stabilité fournit des garanties analytiques sur le comportement des solutions de systèmes contrôlés au fil du temps, qu'elles convergent vers un point ou un ensemble souhaité, et que si elles sont initialement proches d'un attracteur, elles le resteront pour tous les temps futurs. La stabilité assure également une robustesse nominale aux perturbations non modélisées, bien qu'une attention particulière soit nécessaire pour garantir cela pour les systèmes non linéaires en temps discret.

Certains auteurs ont étudié la stabilité dans ce contexte, mais souffrent de certaines ou de multiples limitations, comme 1) la connaissance d'un contrôleur initial de stabilisation globale qui peut être difficile à déterminer ; 2) la stabilité d'un point est étudiée, alors que le système en boucle fermée peut avoir un type d'attracteurs plus général ; 3) nous voulons considérer des coûts d'étape plus généraux, et pas nécessairement quadratiques, ni des fonctions définies positives de l'état et de l'entrée ; 4) ces résultats ne tiennent pas compte

d'un facteur d'actualisation ; 5) et sont sujets à un certain conservatisme car ils fournissent une stabilité asymptotique alors que nous savons que seule une stabilité pratique peut être assurée en général ; 6) reposent sur des hypothèses fortes sur les erreurs d'approximation. Il est donc nécessaire de disposer d'outils génériques pour analyser la stabilité des systèmes non linéaires à temps discret contrôlés par algorithmes de programmation dynamique.

L'objectif de ma thèse est d'étudier la stabilité de systèmes non-linéaires commandés par ces algorithmes. L'enjeu est important car cela permettra de créer un nouveau pont entre l'intelligence artificielle et l'automatique. La stabilité nous informe sur le comportement du système en fonction du temps et garantit sa robustesse en présence de perturbations ou d'incertitudes de modèle. Les algorithmes en intelligence artificielle se concentrent sur l'optimalité de la commande et n'exploitent pas les propriétés de la dynamique du système. La stabilité est non seulement désirable pour les raisons auparavant, mais aussi pour la possibilité de l'exploiter pour améliorer ces algorithmes de intelligence artificielle. Mes travaux de recherche se concentrent sur les techniques de commande issues de la programmation dynamique (approchée) lorsque le modèle du système est connu. J'identifie pour cela des conditions générales grâce auxquelles il est possible de garantir la stabilité du système en boucle fermée. En contrepartie, une fois la stabilité établie, nous pouvons l'exploiter pour améliorer drastiquement les garanties de stabilité de la littérature.

Mes travaux se sont concentrés autour de deux axes. Le premier concerne l'approche par itération sur les valeurs, qui est l'un des piliers de la programmation dynamique approchée et est au coeur de nombre d'algorithmes d'apprentissage par renforcement [SB17]. La fonction de coût considéré est décompté, c'est-à-dire que la fonction à minimiser est pondéré par un 'facteur d'oubli' ou 'de décompte' qui décroît avec le temps. Ce type de fonctions est souvent considéré en programmation dynamique [B12], en apprentissage par renforcement [BBDSE10], et en planification optimiste [LV06] par exemple, et offre de nombreux avantages pour la synthèse et l'analyse d'optimalité. En contrepartie, le

facteur d'oubli est source de difficultés lorsqu'on s'intéresse à la stabilité. A l'évidence, des travaux récents [PBND14, PBND16] démontrent les difficultés techniques de la présence du facteur d'oubli, et que celui-ci doit être choisi suffisamment grand (ergo, un oubli négligeable), mais ces résultats ne sont pas adaptés au cas de l'algorithme d'itération sur les valeurs. Aux apparences, le facteur d'oubli pose un dilemme entre la garantie d'une convergence (rapide) de la stabilité et la garantie de sous optimalité des algorithmes d'intelligence artificielle. Une analyse plus fine et plus proche de l'algorithme est requise. En réalité, les fonctions de coût à horizon infini décompté sont approchées par des fonctions de coût à horizon *fini* décompté. Avec les mêmes hypothèses que [PBND14, PBND16], je démontre la stabilité de contrôle optimal à coût d'horizon fini décompté quand le facteur d'oubli et l'horizon fini sont suffisamment grands. Ces hypothèses sont différentes de celles considérées en programmation dynamique [B12], mais néanmoins suffisantes pour des garanties d'approximation.

De plus, mes résultats se distinguent de ceux existant en programmation dynamique en deux points: ils (i) généralisent au cas non-décompté (sans facteur d'oubli); (ii) s'appliquent à des coûts plus usuels pour des automaticiens. Une (bonne) surprise de cette généralisation est que ces estimations ont moins de conservatismes pour les garanties de stabilité de contrôle optimal à coût non-décompté quand comparées à des travaux précédents [GMTT05], sous certaines conditions et vérifiées dans plusieurs exemples. En suite, j'étends l'analyse au cadre plus général de valeur itération approchée, quand la fonction de coût fini décompté est elle aussi approchée: la stabilité est garantie quand l'approximation est suffisamment bonne. J'applique enfin mes résultats à une simulation de pendule inverse avec commande calculée par valeur itération approchée; celle-ci corrobore mes résultats analytiques. Ces travaux ont été publiés dans [CDC2018]. La version journal est en cours d'évaluation [TAC20XX].

Dans un deuxième temps, je me suis intéressé à l'algorithme de planification optimiste

[LV06], car lui aussi approche la fonction de coût d'horizon infini par un problème d'horizon fini. Cet algorithme de planification optimiste est très puissant pour les systèmes dont l'espace d'entrées est fini, à l'instar des systèmes commutés pour lesquels la loi de commutation est le signal de contrôle. En exploitant la forme du système, l'algorithme calcule exactement la fonction de coût à horizon fini en utilisant intelligemment le budget de calcul disponible. Cependant, ils souffrent des mêmes restrictions que l'étude antérieure: les garanties d'optimalité ne se généralisent pas au cas non-décompté et considèrent des fonctions de coût non-usuelles pour des automaticiens. Je propose donc un algorithme, dénommé OP_{\min} décliné en deux versions, qui ne souffre pas de ces restrictions. En s'appuyant sur les hypothèses suffisantes pour la stabilité, je démontre des garanties d'optimalité, stabilité et de performance pour deux stratégies de contrôle optimal: soit par un budget de calcul restreint mais suffisamment large; soit par une condition d'arrêt de l'algorithme conçu par l'utilisateur. OP_{\min} constitue l'une des premières solutions au contrôle (sous-)optimal, stabilisant des systèmes non-linéaires en temps discrets. OP_{\min} est appliqué dans un cas d'étude pour unification optimale de plusieurs lois de commandes, et aussi dans quelques exemples par simulations. Ces travaux seront présentés en [CDC2019], une version journal est en cours de préparation.

Avec du recul, j'ai pu interpréter de manière concise et intuitive mes résultats précédents en exploitant la notion de ce que j'appelle le "regret". Ici, la notion de regret est liée aux conséquences (coût) des actions "regrettables". Le regret d'une action à un état est défini comme le coût total excédentaire qui résulte de cet apport sous-optimal (ou non). En d'autres termes, les intrants regrettables peuvent avoir un coût supplémentaire "immédiat", un coût "futur" (à faire), ou un mélange des deux. Lorsque le regret d'une action est suffisamment inférieur au coût immédiat, il en résulte de multiples conséquences positives: convergence des coûts, stabilité, garantie des coûts en boucle fermée, qu'il s'agisse du coût moyen, du coût de passage ou du coût de fonctionnement.

J'ai également participé à une étude sur l'analyse de stabilité de systèmes contrôlé par valeur itération approchée, soumis a une conference [CDC20YY]

Cette thèse est encadrée par Jamal Daafouz et Romain Postoyan tous deux membres du département 'Contrôle - Identification - Diagnostic' du CRAN (UMR 7039), et les travaux sont menés en collaboration avec Lucian Buşoniu (Cluj-Napoca Technical University, Roumanie) et Dragan Nešić (The University of Melbourne, Australie).

[SB17] R. S. Sutton and A. G. Barto. "Reinforcement Learning: An Introduction." MIT Press, Cambridge, USA, 2nd edition, 2017

[B12] D. P. Bertsekas, "Dynamic Programming and Optimal Control", volume 2, Athena Scientific, Belmont, 4th edition, Etats-Unis, 2012.

[BBDSE10] L. Buşoniu, R. Babuska, B. De Schutter, and D. Ernst, "Reinforcement Learning and Dynamic Programming Using Function Approximators", Automation and Control Engineering. Taylor & Francis CRC Press, 2010.

[LV06] S. M. LaValle, "Planning Algorithms", Cambridge University Press, New York, Etats-Unis, 2006.

[M14] R. Munos, "The optimistic principle applied to games, optimization and planning: towards foundations of Monte-Carlo tree search", "Foundations and Trends in Machine Learning", 7(1):1- 130, 2014.

[PBND14] R. Postoyan, L. Buşoniu, D. Nešić et J. Daafouz, "Stability of infinite-horizon optimal control with discounted cost", CDC (IEEE Conference on Decision and Control), Los Angeles : Etats- Unis, 2014.

[PBND16] R. Postoyan, L. Buşoniu, D. Nešić et J. Daafouz, "Stability analysis of discrete-time infinite-horizon optimal control with discounted cost", IEEE Transactions on Automatic Control, 2016,

[GMTT05] G. Grimm, M. J. Messina, S. E. Tuna and A. R. Teel, "Model predictive control: for want of a local control Lyapunov function, all is not lost," IEEE Transactions on

Automatic Control, 2005.

[CDC2018] M. Granzotto, R. Postoyan, L. Buşoniu, D. Nešić, J. Daafouz. "Stability analysis of discrete-time finite-horizon discounted optimal control." 57th IEEE Conference on Decision and Control, CDC 2018, Dec 2018, Miami, FL, United States.

[TAC20XX] M. Granzotto, R. Postoyan, L. Buşoniu, D. Nešić, J. Daafouz. "Finite horizon discounted optimal control - stability and performance." Submitted for journal publication at Trans. of Automatic Control at 2/12/2018.

[CDC2019] M. Granzotto, R. Postoyan, L. Buşoniu, D. Nešić, J. Daafouz. "Optimistic planning for the near-optimal control of nonlinear switched discrete-time systems with stability guarantees." Submitted for conference publication at CDC 2019.

[CDC20YY] R. Postoyan, M. Granzotto, L. Buşoniu, B. Scherrer, D. Nešić, J. Daafouz. "Stability guarantees for nonlinear discrete-time systems controlled by approximate value iteration." Submitted for conference publication at CDC 2019.

Mots-clés: Automatique.

