University Library



A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Postoyan, R;Granzotto, M;Busoniu, L;Scherrer, B;Nesic, D;Daafouz, J

Title:
Stability guarantees for nonlinear discrete-time systems controlled by approximate value iteration

Date:
2020-03-12

Citation:
Postoyan, R., Granzotto, M., Busoniu, L., Scherrer, B., Nesic, D. & Daafouz, J. (2020). Stability guarantees for nonlinear discrete-time systems controlled by approximate value iteration. 2019 IEEE 58TH CONFERENCE ON DECISION AND CONTROL (CDC), 2019-December, pp.487-492. IEEE. https://doi.org/10.1109/CDC40024.2019.9030056.

Persistent Link:
https://hdl.handle.net/11343/249534

HAL
archives-ouvertes.fr

Stability guarantees for nonlinear discrete-time systems controlled by approximate value iteration

Romain Postoyan, Mathieu Granzotto, Lucian Buşoniu, Bruno Scherrer, Dragan Nešić, Jamal Daafouz

▶ **To cite this version:**

Romain Postoyan, Mathieu Granzotto, Lucian Buşoniu, Bruno Scherrer, Dragan Nešić, et al.. Stability guarantees for nonlinear discrete-time systems controlled by approximate value iteration. 58th IEEE Conference on Decision and Control, CDC 2019, Dec 2019, Nice, France. hal-02271268

HAL Id: hal-02271268
https://hal.archives-ouvertes.fr/hal-02271268

Submitted on 26 Aug 2019

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.

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, $\mathbb{R}_{\geqslant 0} := [0, \infty)$, $\mathbb{R}_{>0} := (0, \infty)$, $\mathbb{Z}$ be the set of integers, $\mathbb{Z}_{\geqslant 0} := \{0, 1, 2, \ldots\}$ and $\mathbb{Z}_{>0} := \{1, 2, \ldots\}$. We denote by $\mathcal{B}_n(\epsilon)$ the closed ball of $\mathbb{R}^n$ centered at the origin of radius $\epsilon > 0$, $n \in \mathbb{Z}_{>0}$. 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 $\chi : [0, a) \to \mathbb{R}_{\geqslant 0}$ 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 $\mathcal{K}_\infty$ if, in addition, $a = \infty$ and it is unbounded. A continuous function $\chi : [0, a) \times \mathbb{R}_{\geqslant 0} \longrightarrow \mathbb{R}_{\geqslant 0}$ with $a \in \mathbb{R}_{>0} \cup \{\infty\}$ is of class $\mathcal{KL}$ if for each $t \in \mathbb{R}_{\geqslant 0}$, $\chi(\cdot, t)$ 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 $\chi : [0, a) \to \mathbb{R}_{\geqslant 0}$, with $a \in \mathbb{R}_{>0} \cup \{\infty\}$, is of class $\mathcal{N}$ when it is non-decreasing. We say that a function $\chi : [0, a_1) \times [0, a_2) \longrightarrow \mathbb{R}_{\geqslant 0}$, 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 $\mathbb{R}_{\geqslant 0}$ to $\mathbb{R}_{\geqslant 0}$. 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

$$x_{k+1} \quad = \quad f(x_k, u_k), \tag{1}$$

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, $k \in \mathbb{Z}_{\geqslant 0}$ is the time, $f : \mathcal{W} \to \mathbb{R}^{n_x}$ 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

$$J_\gamma(x, \mathbf{u}) \quad := \quad \sum_{k=0}^{\infty} \gamma^k \ell(\phi(k, x, \mathbf{u}|_k), u_k), \tag{2}$$

where $x \in \mathbb{R}^{n_x}$, $\mathbf{u} = (u_0, u_1, \ldots)$ is an infinite-length sequence of admissible inputs, $\gamma \in (0, 1]$ is possibly sub-unitary, $\ell : \mathcal{W} \to \mathbb{R}_{\geqslant 0}$ 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 $\mathbf{u}|_k := (u_0, \ldots, u_{k-1})$, which is the truncation of $\mathbf{u}$ to the first $k \in \mathbb{Z}_{>0}$ steps[1].

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 $\widehat{V}^0 : \mathbb{R}^{n_x} \to \mathbb{R}_{\geqslant 0}$. We then iterate it as follows, for any $x \in \mathbb{R}^{n_x}$ and $i \in \mathbb{Z}_{\geqslant 0}$,

$$\begin{aligned} \widehat{V}^{i+1}_{\gamma, \varepsilon_V}(x) \quad = \quad & \min_{u \in \mathcal{U}(x)} \left[\ell(x, u) + \gamma \widehat{V}^i_{\gamma, \varepsilon_V}(f(x, u)) \right] \\ & + e_{\widehat{V}, i+1}(\gamma, \varepsilon_V, x), \end{aligned} \tag{3}$$

where $\widehat{V}^0_{\gamma, \varepsilon_V} := \widehat{V}^0$ and $e_{\widehat{V}, i+1} : (0, 1] \times \mathcal{B}_{n_V}(\bar{\epsilon}) \times \mathbb{R}^{n_x} \to \mathbb{R}$ is 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 $\widehat{V}^i_{\gamma, \varepsilon_V}$ 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 $\mathcal{B}_{n_V}(\bar{\epsilon})$, where $\bar{\epsilon} \in \mathbb{R}_{>0}$ is an upper-bound on the norm of[2] 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 $\widehat{V}^i_{\gamma, \varepsilon_V}$. The guarantees given in [19] for stochastic problems ensure that

[1] We use the convention $\phi(0, x, \mathbf{u}|_0) = x$ where $\mathbf{u}|_0$ is the empty set.

[2] The adjustable parameter ε_V, later ε, is constrained in $\mathcal{B}_{n_V}(\bar{\epsilon})$, later $\mathcal{B}_{n_\varepsilon}(\bar{\epsilon})$. The forthcoming results straightforwardly apply when we have additional requirements, by constraining accordingly ε_V, ε, 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 $\gamma \in (0,1]$, $\varepsilon_V \in \mathcal{B}_{n_V}(\bar{\epsilon})$, $x \in \mathbb{R}^{n_x}$ and $i \in \mathbb{Z}_{\geqslant 0}$, $\widehat{V}^i_{\gamma,\varepsilon_V}(x) \geqslant 0$. □

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 $\ell(\mathcal{W}) \subseteq \mathbb{R}_{\geqslant 0}$. It can be enforced by suitably selecting the function space where $\widehat{V}^i_{\gamma,\varepsilon_V}$, $i \in \mathbb{Z}_{\geqslant 0}$, is constrained to lie.

The policy, or feedback, calculated at iteration $i \in \mathbb{Z}_{\geqslant 0}$, is given by

$$\begin{aligned}\hat{h}^i_{\gamma,\varepsilon_h}(x) \quad \in \quad & \operatorname{argmin}_{u\in\mathcal{U}(x)} \Big[\ell(x,u) + \gamma \widehat{V}^i_{\gamma,\varepsilon_V}(f(x,u))\Big] \\ & + e_{\hat{h},i}(\gamma, \varepsilon_h, x),\end{aligned} \tag{4}$$

where $e_{\hat{h},i} : (0,1] \times \mathcal{B}_{n_h}(\bar{\epsilon}) \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_u}$ is another approximation error function, which may arise when computing the argmin above, and which can be tuned via parameter $\varepsilon_h \in \mathcal{B}_{n_h}(\bar{\epsilon})$, $n_h \in \mathbb{Z}_{>0}$. The norm of ε_h is subject to the same constraint as ε_V, namely $|\varepsilon_h| \leqslant \bar{\epsilon}$, without loss of generality as we can always define $\bar{\epsilon}$ 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 $\gamma \in (0,1]$, $\varepsilon_h \in \mathcal{B}_{n_h}(\bar{\epsilon})$, $x \in \mathbb{R}^{n_x}$ and $i \in \mathbb{Z}_{\geqslant 0}$, $\hat{h}^i_{\gamma,\varepsilon_h}(x) \in \mathcal{U}(x)$. □

We stop the iterating procedure (3) at any step i bigger than $i^\star \in \mathbb{Z}_{>0}$, which is assumed to satisfy the next property.

Standing Assumption 3 (SA3): There exist $i^\star \in \mathbb{Z}_{\geqslant 0}$ and $e_{\text{stop},i} : (0,1] \times \mathcal{B}_{n_V}(\bar{\epsilon}) \times \mathcal{B}_{n_{\text{stop}}}(\bar{\epsilon}) \times \mathbb{R}^{n_x} \to \mathbb{R}_{\geqslant 0}$ with $n_{\text{stop}} \in \mathbb{Z}_{>0}$, such that for any $\gamma \in (0,1]$, $\varepsilon_V \in \mathcal{B}_{n_V}(\bar{\epsilon})$, $\varepsilon_{\text{stop}} \in \mathcal{B}_{n_{\text{stop}}}(\bar{\epsilon})$, $x \in \mathbb{R}^{n_x}$ and $i \geqslant i^\star$,

$$\mathcal{T}\widehat{V}^i_{\gamma,\varepsilon_V}(x) \quad \leqslant \quad \widehat{V}^i_{\gamma,\varepsilon_V}(x) + e_{\text{stop},i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) \tag{5}$$

where $\mathcal{T}\widehat{V}^i_{\gamma,\varepsilon_V}(x) := \min_{u\in\mathcal{U}(x)} \Big[\ell(x,u) + \gamma \widehat{V}^i_{\gamma,\varepsilon_V}(f(x,u))\Big]$. □

Error $e_{\text{stop},i}$ is called the Bellman residual [28] of $\widehat{V}^i_{\gamma,\varepsilon_V}$, $i \geqslant i^\star$. It is allowed to depend on γ, ε_V, x and also $\varepsilon_{\text{stop}} \in \mathcal{B}_{n_{\text{stop}}}(\bar{\epsilon})$ 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 $\widehat{V}^i_{\gamma,\varepsilon_V}$. It can then easily be shown that the Bellman residual in (5) is at most $(1+\gamma)|\widehat{V}^i_{\gamma,\varepsilon_V}(x) - V^\star_{\gamma,\infty}(x)|$, where $V^\star_{\gamma,\infty}$ is the optimal value function of cost (2). □

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 $\widehat{V}^i_{\gamma,\varepsilon}$ instead of $\widehat{V}^i_{\gamma,\varepsilon_V}$ for the sake of convenience. We similarly write $\hat{h}^i_{\gamma,\varepsilon}$ instead of $\hat{h}^i_{\gamma,\varepsilon_h}$.

As a result, the closed-loop feedback system is given by

$$\begin{aligned}x_{k+1} \in & f\Big(x_k, \operatorname{argmin}_{u_k\in\mathcal{U}(x_k)} \Big[\ell(x_k,u_k) + \gamma \widehat{V}^i_{\gamma,\varepsilon}(f(x_k,u_k))\Big] \\ & + e_{\hat{h},i^\star}(\gamma, \varepsilon_h, x_k)\Big) \\ =: & \widehat{F}_{\gamma,\varepsilon,i}(x_k).\end{aligned} \tag{6}$$

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 $\sigma : \mathbb{R}^{n_x} \to \mathbb{R}_{\geqslant 0}$ 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 $|\cdot|^2_{\mathcal{A}}$ 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 $W : \mathbb{R}^{n_x} \to \mathbb{R}_{\geqslant 0}$, $\alpha_W \in \mathcal{K}_\infty$ and $\overline{\alpha}_W : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ continuous, nondecreasing and zero at zero, such that the following holds for any $(x,u) \in \mathcal{W}$

$$\begin{aligned} W(x) \quad & \leqslant \quad \overline{\alpha}_W(\sigma(x)) \\ W(f(x,u)) - W(x) \quad & \leqslant \quad -\alpha_W(\sigma(x)) + \ell(x,u). \end{aligned} \tag{7}$$

□

Assumption 2 is indeed a detectability property. The best way to see it is when (7) holds with $W = 0$. Then, (7) implies that $\ell(x,u) \geqslant \alpha_W(\sigma(x))$, 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 $W(f(x,u)) - W(x) \leqslant -\alpha_W(\sigma(x)) + \chi(\ell(x,u))$, where $\chi \in \mathcal{K}_\infty$, as in [9, SA3]. This extension is left for future work. □

B. Approximation errors

We make assumptions on the approximation errors $e_{\widehat{V},i}$, $e_{\hat{h},i}$ and $e_{\text{stop},i}$ with $i \geqslant i^\star$ 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 $\mathcal{X}$. □

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 $\hat{h}^i_{\gamma,\varepsilon}(x)$ as in (4). We define

$$h^i_{\gamma,\varepsilon}(x) \quad := \quad \hat{h}^i_{\gamma,\varepsilon}(x) - e_{\hat{h},i}(\gamma,\varepsilon_h,x). \tag{8}$$

which corresponds to the feedback $\hat{h}^i_{\gamma,\varepsilon}$ in (4) not affected by the error term $e_{\hat{h},i}$. Hence $h^i_{\gamma,\varepsilon}(x) \in \operatorname{argmin}_{u\in\mathcal{U}(x)} \left[\ell(x,u) + \gamma \widehat{V}^i_{\gamma,\varepsilon_V}(f(x,u))\right]$. We make the next assumption.

Assumption 4: There exist $\alpha_{\text{stop}}, \widehat{\alpha}_V, \widehat{\alpha}_\ell \in \mathcal{KN}$, $\underline{\alpha}_W, \underline{\alpha}_\varepsilon \in \mathcal{K}$ such that, for any $\gamma \in (0,1]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, $x \in \mathcal{X}$ and $i \geqslant i^\star$,

$$e_{\text{stop},i}(\gamma,\varepsilon_V,\varepsilon_{\text{stop}},x) \leqslant \alpha_{\text{stop}}(|\varepsilon|,\sigma(x)) \tag{9a}$$

$$\begin{aligned}\min\Big\{|e_{\widehat{V},i}(\gamma,\varepsilon_V,x)|, e_{\text{stop},i}(\gamma,\varepsilon_V,\varepsilon_{\text{stop}},x)\Big\}& \\ \leqslant \alpha_W(\sigma(x)) - \underline{\alpha}_W(\sigma(x)) + \underline{\alpha}_\varepsilon(|\varepsilon|)&\end{aligned} \tag{9b}$$

$$\begin{aligned}\left|\widehat{V}^i_{\gamma,\varepsilon}(f(x,h^i_{\gamma,\varepsilon}(x))) - \widehat{V}^i_{\gamma,\varepsilon}(f(x,\hat{h}^i_{\gamma,\varepsilon}(x)))\right|& \\ \leqslant \widehat{\alpha}_V(|\varepsilon|,\sigma(x))&\end{aligned} \tag{9c}$$

$$\left|\ell(x,h^i_{\gamma,\varepsilon}(x)) - \ell(x,\hat{h}^i_{\gamma,\varepsilon}(x))\right| \leqslant \widehat{\alpha}_\ell(|\varepsilon|,\sigma(x)). \tag{9d}$$

□

Condition (9a) implies that the iteration (3) practically converges, so that the mismatch $\mathcal{T}\widehat{V}^i_{\gamma,\varepsilon_V} - \widehat{V}^i_{\gamma,\varepsilon_V}$ 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, $\varepsilon = \epsilon_i\bar{\varepsilon}$ 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 $|e_{\widehat{V},i}(\gamma,\varepsilon_V,x)| \leqslant \overline{\alpha}_{V,\varepsilon}(|\varepsilon|)$ with $\overline{\alpha}_{V,\varepsilon} \in \mathcal{K}$ 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 $\underline{\alpha}_\varepsilon = \overline{\alpha}_{V,\varepsilon}$. 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 $\widehat{V}^i_{\gamma,\varepsilon}(f(x,\cdot))$ for $i \geqslant i^\star$ evaluated at the applied policy $\hat{h}^i_{\gamma,\varepsilon_h}(x)$ and at the ideal one $h^i_{\gamma,\varepsilon_h}(x)$ 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 $i \geqslant i^\star$, $\widehat{V}^i_{\gamma,\varepsilon}$ is continuous on $\mathcal{X}$.

(ii) There exists $\alpha_\sigma \in \mathcal{K}_\infty$ such that $\sigma(x) \geqslant \alpha_\sigma(|x|)$ 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 $|\ell(x,u) - \ell(x,v)| \leqslant \theta_\ell(|u-v|)$ 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 $\vartheta_h : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ non-decreasing such that, for any $\gamma \in (0,1]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, $i \geqslant i^\star$, $x \in \mathcal{X}$, $\sigma(x) \geqslant \alpha_\sigma(|x|)$ and $|h^i_{\gamma,\varepsilon}(x)| \leqslant \vartheta_h(\sigma(x))$.

(iii) There exist $\alpha_h \in \mathcal{KN}$ such that, for any $\gamma \in (0,1]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, $i \geqslant i^\star$ and $x \in \mathcal{X}$, $|e_{\hat{h},i}(\gamma,\varepsilon_h,x)| \leqslant \alpha_h(|\varepsilon_h|,\sigma(x))$.

Then (9c)-(9d) hold. □

The continuity of $\widehat{V}^i_{\gamma,\varepsilon}$ in item (i) of Lemma 1 can be enforced by suitably selecting the function space where $\widehat{V}^i_{\gamma,\varepsilon}$ 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 $h^i_{\gamma,\varepsilon}$, 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 $\widehat{V}^i_{\gamma,\varepsilon}$ with $i \geqslant i^\star$.

Assumption 5: There exist $\overline{\alpha}_V, \overline{\alpha}_\varepsilon \in \mathcal{K}_\infty$ such that for any $\gamma \in (0,1]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, $i \geqslant i^\star$, $x \in \mathcal{X}$,

$$\widehat{V}^i_{\gamma,\varepsilon}(x) \quad \leqslant \quad \overline{\alpha}_V(\sigma(x)) + \overline{\alpha}_\varepsilon(|\varepsilon|). \tag{10}$$

□

Assumption 5 is related to the stabilizability of system (1). To see it, assume there exists $\breve{\alpha}_V \in \mathcal{KN}$ such that for any $\gamma \in (0,1]$, $\varepsilon_V \in \mathcal{B}_{n_V}(\bar{\varepsilon})$, $x \in \mathcal{X}$ and $i \in \mathbb{Z}_{\geqslant 0}$, $|e_{\widehat{V},i}(\gamma,\varepsilon_V,x)| \leqslant \breve{\alpha}_V(|\varepsilon_V|,\sigma(x))$. Such an assumption is generally needed for the satisfaction of the first two inequalities in (9). Define the modified stage cost $\widetilde{\ell}(x,u) := \ell(x,u) + \breve{\alpha}_V(\bar{\varepsilon},\sigma(x))$ for any $(x,u) \in \mathcal{W}$. The associated *undiscounted* optimal value function is $\widetilde{V}(x) := \min_{u\in\mathcal{U}(x)} \sum_{k=0}^{\infty} \widetilde{\ell}(\phi(k,x,\mathbf{u}|_k),u_k)$ for any $x \in \mathcal{X}$. The next lemma gives a condition on $\widetilde{V}$ ensuring Assumption 5.

Lemma 2: Under Assumption 4, if there exists $\widetilde{\alpha}_V \in \mathcal{K}_\infty$ such that, for any $x \in \mathcal{X}$,

$$\widetilde{V}(x) \quad \leqslant \quad \widetilde{\alpha}_V(\sigma(x)) \tag{11}$$

then Assumption 5 holds with $\overline{\alpha}_V = \widetilde{\alpha}_V$ and $\overline{\alpha}_\varepsilon = 0$. □

Proof. The proof is similar to [10, Section V.B]. In view of the definition of $\widetilde{\ell}$ and Assumption 4, $\gamma\ell(x,u) \leqslant \ell(x,u) \leqslant \widetilde{\ell}(x,u)$ for any $(x,u) \in \mathcal{W}$ and $\gamma \in (0,1]$. Hence, by definition of $\widetilde{V}$, for any $\gamma \in (0,1]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\epsilon})$, $x \in \mathcal{X}$, $i \geqslant i^\star$, it holds $\widehat{V}^i_{\gamma,\varepsilon}(x) \leqslant \widetilde{V}(x)$. The desired result is obtained by invoking (11). ■

Condition (11) is a stabilizability property of system (1) with respect to stage cost $\widetilde{\ell}$ according to [9, Section III] and in [21, Lemma 1], which we recall below.

Lemma 3: Consider system (1) and suppose that $\widetilde{\ell}$ is *uniformly globally exponentially stabilizable to zero with respect to* $\sigma : \mathbb{R}^{n_x} \to \mathbb{R}_{\geqslant 0}$, 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 $\widetilde{\ell}(\phi(k,x,\mathbf{u}|_k(x)), u_k(x)) \leqslant M\sigma(x)e^{-\lambda k}$ for any $k \in \mathbb{Z}_{>0}$. Then (11) holds with $\widetilde{\alpha}_V(s) = \frac{Ms}{1-e^{-\lambda}}$ for any $s \geqslant 0$. □

The works in [17], [26] require (10) to hold with $\overline{\alpha}_\varepsilon = 0$, see the proof of Theorem 4 in [26] for instance. On the other hand, condition (11) typically requires $e_{\widehat{V},i+1}(\gamma, \varepsilon_V, x)$ 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 $\overline{\alpha}_\varepsilon \neq 0$ 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. □

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 $\alpha_Y, \underline{\alpha}_Y, \overline{\alpha}_Y \in \mathcal{K}_\infty$ and $\Upsilon \in \mathcal{KN}$ such that for any $(\gamma,\varepsilon) \in (0,1] \times \mathcal{B}_{n_\varepsilon}(\bar{\epsilon})$, and $i \geqslant i^\star$ where $i^\star$ comes from SA3, function $Y^i_{\gamma,\varepsilon} := \widehat{V}^i_{\gamma,\varepsilon} + W$ defined on $\mathbb{R}^{n_x}$ satisfies the following.

(a) For any $x \in \mathcal{X}$, $\underline{\alpha}_Y(\sigma(x)) - \underline{\alpha}_\varepsilon(|\varepsilon|) \leqslant Y^i_{\gamma,\varepsilon}(x) \leqslant \overline{\alpha}_Y(\sigma(x)) + \overline{\alpha}_\varepsilon(|\varepsilon|)$.

(b) For any $x \in \mathcal{X}$, $\upsilon \in \widehat{F}_{\gamma,\varepsilon,i}(x)$, $Y^i_{\gamma,\varepsilon}(\upsilon) - Y^i_{\gamma,\varepsilon}(x) \leqslant -\alpha_Y(\sigma(x)) + \Upsilon\big(\frac{1-\gamma}{\gamma} + |\varepsilon|, \sigma(x)\big)$, where $\widehat{F}_{\gamma,\varepsilon,i}$ is defined in (6).

The expressions of $\underline{\alpha}_Y$, $\overline{\alpha}_Y$, α_Y and Υ are provided in Table I. □

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 $(\gamma^\star, \epsilon^\star, \Delta) \in (0,1) \times (0,\bar{\epsilon}) \times \mathbb{R}_{>0}$ such that for any $i \geqslant i^\star$, $\gamma \in [\gamma^\star, 1]$, $\epsilon \in [0, \epsilon^\star]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\epsilon)$, $x \in \mathcal{X}$ with $Y^i_{\gamma,\varepsilon}(x) \leqslant \Delta$, where $Y^i_{\gamma,\varepsilon}$ is defined in Theorem 1, any solution[3] $\phi(\cdot, x)$ to system (6) satisfies $\phi(k,x) \in \mathcal{X}$ and

$$\sigma(\phi(k,x)) \quad \leqslant \quad \beta(\sigma(x), k) + \delta, \tag{12}$$

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

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 $Y^i_{\gamma,\varepsilon}(x) \leqslant \Delta$ 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 $\gamma^\star$, the maximal error $\epsilon^\star$, 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. □

B. Stronger statements

The next result ensures an asymptotic and exponential stability properties for system (6) and the bounds $\gamma^\star$ and $\epsilon^\star$ are independent of constant Δ, contrary to[4] Theorem 2.

Corollary 1: Consider system (6) and suppose the following holds.

(i) Assumptions 1 and 3 hold.

(ii) $e_{\hat{h},i} = 0$ in (4) for any $i \in \mathbb{Z}_{\geqslant 0}$.

[3] We use the same notation to denote a solution to (6) and to (1), see (2), with some slight abuse.

[4] See (24).

$\underline{\alpha}_Y :=$	$\underline{\alpha}_W$	
$\overline{\alpha}_Y :=$	$\overline{\alpha}_V + \overline{\alpha}_W$	
$\alpha_Y :=$	α_W	
$\Upsilon :=$	$(s_1, s_2) \mapsto \frac{1-\gamma}{\gamma}\Big(\overline{\alpha}_V(s_2) + \overline{\alpha}_\varepsilon(s_1) + \alpha_{\text{stop}}(s_1, s_2)\Big) + \hat{\alpha}_\ell(s_1, s_2) + \alpha_{\text{stop}}(s_1, s_2) + \hat{\alpha}_V(s_1, s_2)$	

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 $s, s_1, s_2 \geqslant 0$.

Let $(\gamma^\star, \epsilon^\star) \in (0,1) \times (0, \bar{\epsilon})$ be such that

$$\frac{1-\gamma^\star}{\gamma^\star}\Big(\bar{a}_V + a_{\text{stop}}(\epsilon^\star)\Big) + a_{\text{stop}}(\epsilon^\star) \quad < \quad a_W. \tag{13}$$

There exist $\beta \in \mathcal{KL}$ and $\Delta > 0$ such that for any $i \geqslant i^\star$, $\gamma \in [\gamma^\star, 1]$, $\epsilon \in [0, \epsilon^\star]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\epsilon)$, and $x \in \mathcal{X}$ such that $Y^i_{\gamma,\varepsilon}(z) \leqslant \Delta$, any solution $\phi(\cdot, x)$ to system (6) initialized at $x \in \mathcal{X}$ satisfies $\phi(k, x) \in \mathcal{X}$ and

$$\sigma(\phi(k, x)) \quad \leqslant \quad \beta(\sigma(x), k), \tag{14}$$

for all $k \in \mathbb{Z}_{\geqslant 0}$. In addition, when $\underline{\alpha}_W(s) = \underline{a}_W \cdot s$ for any $s \geqslant 0$, $\beta \in \exp -\mathcal{KL}$. □

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 $(\gamma^\star, \epsilon^\star)$ is independent of Δ, contrary to Theorem 2. Note that condition (13) is always verified for $\gamma^\star$ sufficiently close to 1 and $\epsilon^\star$ 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 $\varepsilon_V \in \mathcal{B}_{n_V}(\bar{\epsilon})$ $|e_{\widehat{V},i+1}(1, \varepsilon_V, x)| \leqslant |\varepsilon_V| \ell(x, 0)$ for any $x \in \mathcal{X}$.
(v) For any $\varepsilon_V \in [0, \bar{\epsilon})$, $x \in \mathbb{R}^{n_x} \backslash \mathcal{X}$ and $i \in \mathbb{Z}_{\geqslant 0}$, $\widehat{V}^i_{\gamma,\varepsilon_V}(x) \geqslant 0$.
(vi) For any $i \in \mathbb{Z}_{\geqslant 0}$, $\varepsilon_h \in \mathbb{R}^{n_h}$ and $x \in \mathbb{R}^{n_x}$, $e_{\hat{h},i}(1, \varepsilon_h, x) = 0$ in (4).
(vii) There exists $d > 0$ such that $\widehat{V}^0(x) \leqslant dQ(x)$ for any $x \in \mathcal{X}$.

Then the conditions of Corollary 1 hold with $\sigma = Q$, $W = 0$, $\overline{\alpha}_W = \alpha_W = \mathbb{I}$, $a_W = 1$, $\overline{\alpha}_V = 2d\mathbb{I}$, $\overline{a}_V = 2d$, $\overline{\alpha}_\varepsilon = 0$, $\alpha_{\text{stop}}(s_1, s_2) = a_{\text{stop}}(s_1)ds_2$ with $a_{\text{stop}}(s_1) = \frac{4s_1}{1-s_1}$ for any $s_1 \in [0, 1)$ and $s_2 \geqslant 0$, $\underline{\alpha}_\varepsilon = 0$, $\underline{\alpha}_W = (1-\bar{\epsilon})\mathbb{I}$ and $\underline{a}_W = 1 - \bar{\epsilon}$. □

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 $\widehat{\mathcal{B}}^i_r$ 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 $\widehat{\mathcal{B}}^i_r$ 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 $\underline{\alpha}_Q(|x|) \leqslant Q(x) \leqslant \overline{\alpha}_Q(|x|)$ for any $x \in \mathcal{X}$, as $\mathcal{X}$ is compact here, in view of[5] [15, Lemma 4.3]. Hence, (14) leads to $|\phi(k, x)| \leqslant \tilde{\beta}(|x|, k)$ with $\tilde{\beta}(s_1, s_2) \mapsto \underline{\alpha}_Q^{-1}\big(\beta(\overline{\alpha}_Q(s_1), s_2)\big) \in \mathcal{KL}$, which is the standard characterization of (local) asymptotic stability of the origin. Moreover, condition (13) is, in this case, $\epsilon^\star < \dfrac{1}{4d+1}$, which is less restrictive than the corresponding condition, namely $\epsilon^\star < 1 + 2d - \sqrt{4d^2 + 4d}$, 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

[5]Lemma 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 $\mathcal{K}_\infty$ 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 $\gamma \in (0,1]$, $\varepsilon \in \mathcal{B}_{n_\varepsilon}(\bar{\varepsilon})$, $i \geqslant i^\star$, $x \in \mathcal{X}$ and $\upsilon \in \widehat{F}_{\gamma,\varepsilon,i}(x)$. Hence, $\upsilon = f(x, \hat{h}^i_{\gamma,\varepsilon}(x))$ for some $\hat{h}^i_{\gamma,\varepsilon}(x)$ defined as in (4). We first prove item (i) of Theorem 1. In view of Assumptions 2 and 5 and the definition of $Y^i_{\gamma,\varepsilon}$, $Y^i_{\gamma,\varepsilon}(x) \leqslant \overline{\alpha}_V(\sigma(x)) + \overline{\alpha}_\varepsilon(|\varepsilon|) + \overline{\alpha}_W(\sigma(x))$. Hence $Y^i_{\gamma,\varepsilon}(x) \leqslant \overline{\alpha}_Y(\sigma(x)) + \overline{\alpha}_\varepsilon(|\varepsilon|)$ where $\overline{\alpha}_Y := \overline{\alpha}_V + \overline{\alpha}_W \in \mathcal{K}_\infty$. In view of SA3,

$$\begin{aligned}\widehat{V}^i_{\gamma,\varepsilon}(x) \;\geqslant\; & \ell(x, h^i_{\gamma,\varepsilon}(x)) + \gamma \widehat{V}^i_{\gamma,\varepsilon}(f(x, h^i_{\gamma,\varepsilon}(x))) \\ & - e_{\text{stop},i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x),\end{aligned} \tag{15}$$

where $h^i_{\gamma,\varepsilon}(x)$ is as in (8). We deduce that $Y^i_{\gamma,\varepsilon}(x) = \widehat{V}^i_{\gamma,\varepsilon}(x) + W(x) \geqslant \ell(x, h^i_{\gamma,\varepsilon}(x)) + \gamma \widehat{V}^i_{\gamma,\varepsilon}(f(x, h^i_{\gamma,\varepsilon}(x))) - e_{\text{stop},i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) + W(x)$. According to SA1, $\gamma \widehat{V}^i_{\gamma,\varepsilon}(f(x, h^i_{\gamma,\varepsilon}(x))) \geqslant 0$, thus $Y^i_{\gamma,\varepsilon}(x) \geqslant \ell(x, h^i_{\gamma,\varepsilon}(x)) - e_{\text{stop},i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) + W(x)$. On the other hand, according to Assumption 2, $W(x) \geqslant \alpha_W(\sigma(x)) - \ell(x, h^i_{\gamma,\varepsilon}(x))$. We derive $Y^i_{\gamma,\varepsilon}(x) \geqslant -e_{\text{stop},i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) + \alpha_W(\sigma(x))$. In view of the last inequality in (9), we have, when $e_{\text{stop},i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) \leqslant |e_{\widehat{V},i}(\gamma, \varepsilon_V, x)|$, $Y^i_{\gamma,\varepsilon}(x) \geqslant \underline{\alpha}_W(\sigma(x)) - \underline{\alpha}_\varepsilon(|\varepsilon|)$. If $e_{\text{stop},i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) > |e_{\widehat{V},i}(\gamma, \varepsilon_V, x)|$, in view of (3),

$$\begin{aligned}\widehat{V}^i_{\gamma,\varepsilon}(x) \;=\; & \ell(x, h^{i-1}_\gamma(x)) + \gamma \widehat{V}^{i-1}_{\gamma,\varepsilon}(f(x, h^{i-1}_\gamma(x))) \\ & + e_{\widehat{V},i}(\gamma, \varepsilon_V, x).\end{aligned} \tag{16}$$

By following similar lines as above, we derive that $Y^i_{\gamma,\varepsilon}(x) \geqslant e_{\widehat{V},i}(\gamma, \varepsilon_V, x) + \alpha_W(\sigma(x)) \geqslant -|e_{\widehat{V},i}(\gamma, \varepsilon_V, x)| + \alpha_W(\sigma(x))$. Then, in view of the last inequality in (9), we derive $Y^i_{\gamma,\varepsilon}(x) \geqslant \underline{\alpha}_W(\sigma(x)) - \underline{\alpha}_\varepsilon(|\varepsilon|)$. 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

$$\begin{aligned}Y^i_{\gamma,\varepsilon}(\upsilon) - Y^i_{\gamma,\varepsilon}(x) \;=\; & \widehat{V}^i_{\gamma,\varepsilon}(\upsilon) - \widehat{V}^i_{\gamma,\varepsilon}(x) \\ & + W(\upsilon) - W(x).\end{aligned} \tag{17}$$

In view of (15),

$$\begin{aligned}Y^i_{\gamma,\varepsilon}(\upsilon) - Y^i_{\gamma,\varepsilon}(x) \;\leqslant\; & \widehat{V}^i_{\gamma,\varepsilon}(\upsilon) - \ell(x, h^i_{\gamma,\varepsilon}(x)) \\ & - \gamma \widehat{V}^i_{\gamma,\varepsilon}(f(x, h^i_{\gamma,\varepsilon}(x))) \\ & + e_{\text{stop},i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) \\ & + W(\upsilon) - W(x).\end{aligned} \tag{18}$$

We obtain, by adding and subtracting $\widehat{V}^i_{\gamma,\varepsilon}(f(x, h^i_{\gamma,\varepsilon}(x)))$,

$$\begin{aligned}Y^i_{\gamma,\varepsilon}(\upsilon) - Y^i_{\gamma,\varepsilon}(x) \;\leqslant\; & \widehat{V}^i_{\gamma,\varepsilon}(\upsilon) - \widehat{V}^i_{\gamma,\varepsilon}(f(x, h^i_{\gamma,\varepsilon}(x))) \\ & - \ell(x, h^i_{\gamma,\varepsilon}(x)) \\ & + (1-\gamma) \widehat{V}^i_{\gamma,\varepsilon}(f(x, h^i_{\gamma,\varepsilon}(x))) \\ & + e_{\text{stop},i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) \\ & + W(\upsilon) - W(x).\end{aligned} \tag{19}$$

From Assumption 2 and by definition of υ, we have

$$W(\upsilon) - W(x) \;\leqslant\; -\alpha_W(\sigma(x)) + \ell(x, \hat{h}^i_{\gamma,\varepsilon}(x)). \tag{20}$$

Therefore, in view of (19),

$$\begin{aligned}Y^i_{\gamma,\varepsilon}(\upsilon) - Y^i_{\gamma,\varepsilon}(x) \leqslant & -\alpha_W(\sigma(x)) \\ & + \ell(x, \hat{h}^i_{\gamma,\varepsilon}(x)) - \ell(x, h^i_{\gamma,\varepsilon}(x)) \\ & + (1-\gamma) \widehat{V}^i_{\gamma,\varepsilon}(f(x, h^i_{\gamma,\varepsilon}(x))) \\ & + e_{\text{stop},i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x) \\ & + \widehat{V}^i_{\gamma,\varepsilon}(\upsilon) - \widehat{V}^i_{\gamma,\varepsilon}(f(x, h^i_{\gamma,\varepsilon}(x))).\end{aligned} \tag{21}$$

The rest of the proof consists in suitably upper-bounding $\ell(x, \hat{h}^i_{\gamma,\varepsilon}(x)) - \ell(x, h^i_{\gamma,\varepsilon}(x))$, $(1-\gamma)\widehat{V}^i_{\gamma,\varepsilon}(f(x, h^i_{\gamma,\varepsilon}(x)))$, $\widehat{V}^i_{\gamma,\varepsilon}(\upsilon) - \widehat{V}^i_{\gamma,\varepsilon}(f(x, h^i_{\gamma,\varepsilon}(x)))$ 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,

$$\widehat{V}^i_{\gamma,\varepsilon}(f(x, h^i_{\gamma,\varepsilon}(x))) \leqslant \frac{1}{\gamma}\left(\widehat{V}^i_{\gamma,\varepsilon}(x) + e_{\text{stop},i}(\gamma, \varepsilon_V, \varepsilon_{\text{stop}}, x)\right). \tag{22}$$

Hence, in view of Assumptions 4 and 5,

$$\begin{aligned}&(1-\gamma)\widehat{V}^i_{\gamma,\varepsilon}(f(x, h^i_{\gamma,\varepsilon}(x))) \\ &\quad \leqslant \frac{1-\gamma}{\gamma}\Big(\overline{\alpha}_V(\sigma(x)) + \overline{\alpha}_\varepsilon(|\varepsilon|) + \alpha_{\text{stop}}(|\varepsilon|, \sigma(x))\Big).\end{aligned} \tag{23}$$

Consequently, by applying Assumption 4, $Y^i_{\gamma,\varepsilon}(\upsilon) - Y^i_{\gamma,\varepsilon}(x) \leqslant -\alpha_W(\sigma(x)) + \hat{\alpha}_\ell(|\varepsilon|, \sigma(x)) + \frac{1-\gamma}{\gamma}\Big(\overline{\alpha}_V(\sigma(x)) + \overline{\alpha}_\varepsilon(|\varepsilon|) + \alpha_{\text{stop}}(|\varepsilon|, \sigma(x))\Big) + \alpha_{\text{stop}}(|\varepsilon|, \sigma(x)) + \hat{\alpha}_V(|\varepsilon|, \sigma(x)) = -\alpha_Y(\sigma(x)) + \Upsilon\Big(\frac{1-\gamma}{\gamma} + |\varepsilon|, \sigma(x)\Big)$, with α_Y and Υ defined as in Table I, which are indeed of class $\mathcal{K}_\infty$ and $\mathcal{KN}$, respectively, in view of the properties of the involved functions. We have proved that item (ii) of Theorem 1 holds. ■

Proof of Theorem 2. Let $\delta > 0$, $i \geqslant i^\star$ and $\tilde{\epsilon} \in (0, \bar{\epsilon}]$ be small enough such that the set $\{x \in \mathbb{R}^{n_x} : \underline{\alpha}_Y(\sigma(x)) \leqslant \underline{\alpha}_\varepsilon(\bar{\varepsilon})\}$ 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 $\tilde{\epsilon}$ 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 $\{x \in \mathbb{R}^{n_x} : \underline{\alpha}_Y(\sigma(x)) \leqslant \underline{\alpha}_\varepsilon(\tilde{\epsilon}) + \Delta\} \subseteq \mathcal{X}$, which exists in view of the definition of $\tilde{\epsilon}$ above, Assumption 1, and the fact that $\mathcal{X}$ is compact. We also define $\tilde{\delta} := \min\left\{\left(\mathbb{I} - \frac{\tilde{\alpha}_Y}{2}\right)^{-1} \circ \underline{\alpha}_Y(\delta), \frac{1}{2}\underline{\alpha}_Y(\frac{1}{2}\delta), \Delta\right\}$ where $\tilde{\alpha}_Y(s) := \alpha_Y \circ \overline{\alpha}_Y^{-1}(\frac{1}{2}s)$ for any $s \geqslant 0$. Note that

$\left(\mathbb{I}-\frac{\widetilde{\alpha}_Y}{2}\right)^{-1}$ is well-defined as we can assume without loss of generality that $\mathbb{I}-\frac{\widetilde{\alpha}_Y}{2}\in\mathcal{K}_\infty$, see [8, footnote 5].

We select $\gamma^\star\in(0,1)$ and $\epsilon^\star\in(0,\bar{\epsilon})$ such that

$$\begin{cases} \Upsilon\left(\frac{1-\gamma^\star}{\gamma^\star}+\epsilon^\star,\underline{\alpha}_Y^{-1}\left(s+\underline{\alpha}_\varepsilon(\epsilon^\star)\right)\right) \leqslant \frac{1}{2}\widetilde{\alpha}_Y(s) \\ \qquad\qquad\qquad\qquad \forall s\in[\tilde{\delta},\Delta] \\ \underline{\alpha}_Y^{-1}\left(2\underline{\alpha}_\varepsilon(\epsilon^\star)\right)+\underline{\alpha}_Y^{-1}\left(2\beta_Y(4\overline{\alpha}_\varepsilon(\epsilon^\star),0)\right) \leqslant \frac{1}{2}\delta \\ 2\overline{\alpha}_\varepsilon(\epsilon^\star)+\Upsilon\left(\frac{1-\gamma^\star}{\gamma^\star}+\epsilon^\star,\overline{\alpha}_Y^{-1}\circ\overline{\alpha}_\varepsilon(\epsilon^\star)\right) \leqslant \tilde{\delta}, \end{cases} \tag{24}$$

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 $(\gamma^\star,\epsilon^\star)\in(0,1)\times(0,\bar{\epsilon})$ such that $\Upsilon\left(\frac{1-\gamma^\star}{\gamma^\star}+\epsilon^\star,\underline{\alpha}_Y^{-1}\left(\Delta+\underline{\alpha}_\varepsilon(\epsilon^\star)\right)\right)\leqslant\frac{1}{2}\widetilde{\alpha}_Y(\tilde{\delta})$, which ensures the first inequality in (24). The second and third inequalities are satisfied by selecting $\epsilon^\star>0$ sufficiently small and large $\gamma^\star\in(0,1)$ in view of the properties of the involved functions.

Let $\gamma\in[\gamma^\star,1]$, $\epsilon\in[0,\epsilon^\star]$, $\varepsilon\in\mathcal{B}_{n_\varepsilon}(\epsilon)$, and $x\in\mathcal{X}$ be[6] such that $Y^i_{\gamma,\varepsilon}(x)\leqslant\Delta$, and $\upsilon\in\widehat{F}_{\gamma,\varepsilon,i}(x)$. We first consider the case where $\overline{\alpha}_Y(\sigma(x))\geqslant\overline{\alpha}_\varepsilon(|\varepsilon|)$. As a result, in view of item (i) of Theorem 1, $Y^i_{\gamma,\varepsilon}(x)\leqslant\overline{\alpha}_Y(\sigma(x))+\overline{\alpha}_\varepsilon(|\varepsilon|)\leqslant 2\overline{\alpha}_Y(\sigma(x))$. In view of Theorem 1, the definition of $\widetilde{\alpha}_Y$ and the fact that $\Upsilon\in\mathcal{KN}$,

$$\begin{aligned} Y^i_{\gamma,\varepsilon}(\upsilon)-Y^i_{\gamma,\varepsilon}(x) \leqslant & -\widetilde{\alpha}_Y(Y^i_{\gamma,\varepsilon}(x)) \\ & +\Upsilon\left(\tfrac{1-\gamma}{\gamma}+|\varepsilon|,\underline{\alpha}_Y^{-1}\left(Y^i_{\gamma,\varepsilon}(x)+\underline{\alpha}_\varepsilon(|\varepsilon|)\right)\right). \end{aligned} \tag{25}$$

Since $\gamma\in[\gamma^\star,1]$, $\varepsilon\in\mathcal{B}_{n_\varepsilon}(\epsilon^\star)$ and $\Upsilon\in\mathcal{KN}$,

$$\begin{aligned} Y^i_{\gamma,\varepsilon}(\upsilon)-Y^i_{\gamma,\varepsilon}(x) \leqslant & -\widetilde{\alpha}_Y(Y^i_{\gamma,\varepsilon}(x)) \\ & +\Upsilon\left(\tfrac{1-\gamma^\star}{\gamma^\star}+\epsilon^\star,\underline{\alpha}_Y^{-1}\left(Y^i_{\gamma,\varepsilon}(x)+\underline{\alpha}_\varepsilon(\epsilon^\star)\right)\right). \end{aligned} \tag{26}$$

When $Y^i_{\gamma,\varepsilon}(x)\in[\tilde{\delta},\Delta]$, in view of (24) and (26),

$$Y^i_{\gamma,\varepsilon}(\upsilon)-Y^i_{\gamma,\varepsilon}(x) \leqslant -\tfrac{1}{2}\widetilde{\alpha}_Y(Y^i_{\gamma,\varepsilon}(x)). \tag{27}$$

When $Y^i_{\gamma,\varepsilon}(x)\in[0,\tilde{\delta}]$, in view of (24), (26), the definition of $\tilde{\delta}$ and since $\mathbb{I}-\widetilde{\alpha}_Y$ can be assumed to be in $\mathcal{K}_\infty$ without loss of generality [8, Remark 5],

$$\begin{aligned} Y^i_{\gamma,\varepsilon}(\upsilon)\leqslant & (\mathbb{I}-\widetilde{\alpha}_Y)\left(Y^i_{\gamma,\varepsilon}(x)\right) \\ & +\Upsilon\left(\tfrac{1-\gamma^\star}{\gamma^\star}+\epsilon^\star,\underline{\alpha}_Y^{-1}\left(Y^i_{\gamma,\varepsilon}(x)+\underline{\alpha}_\varepsilon(\epsilon^\star)\right)\right) \\ \leqslant & (\mathbb{I}-\widetilde{\alpha}_Y)(\tilde{\delta})+\Upsilon\left(\tfrac{1-\gamma^\star}{\gamma^\star}+\epsilon^\star,\underline{\alpha}_Y^{-1}(\tilde{\delta}+\underline{\alpha}_\varepsilon(\epsilon^\star))\right) \\ \leqslant & (\mathbb{I}-\widetilde{\alpha}_Y)(\tilde{\delta})+\tfrac{1}{2}\widetilde{\alpha}_Y(\tilde{\delta}) \\ = & \left(\mathbb{I}-\tfrac{1}{2}\widetilde{\alpha}_Y\right)(\tilde{\delta}). \end{aligned} \tag{28}$$

Thus, whenever $Y^i_{\gamma,\varepsilon}(x)\leqslant\tilde{\delta}$, $Y^i_{\gamma,\varepsilon}(\upsilon)\leqslant\tilde{\delta}$.

We now consider the case where $\overline{\alpha}_Y(\sigma(x))\leqslant\overline{\alpha}_\varepsilon(|\varepsilon|)$, which implies $Y^i_{\gamma,\varepsilon}(x)\leqslant 2\overline{\alpha}_\varepsilon(|\varepsilon|)$ according to item (i) of Theorem 1. Consequently, in view of item (ii) of Theorem 1,

$$\begin{aligned} Y^i_{\gamma,\varepsilon}(\upsilon) \leqslant & Y^i_{\gamma,\varepsilon}(x)-\alpha_Y(\sigma(x)) \\ & +\Upsilon\left(\tfrac{1-\gamma}{\gamma}+|\varepsilon|,\sigma(x)\right) \\ \leqslant & Y^i_{\gamma,\varepsilon}(x)+\Upsilon\left(\tfrac{1-\gamma}{\gamma}+|\varepsilon|,\sigma(x)\right) \\ \leqslant & 2\overline{\alpha}_\varepsilon(|\varepsilon|) \\ & +\Upsilon\left(\tfrac{1-\gamma}{\gamma}+|\varepsilon|,\overline{\alpha}_Y^{-1}\circ\overline{\alpha}_\varepsilon(|\varepsilon|)\right). \end{aligned} \tag{29}$$

Since $|\varepsilon|\leqslant\epsilon^\star$ and $\gamma\in[\gamma^\star,1]$,

$$\begin{aligned} Y^i_{\gamma,\varepsilon}(\upsilon) \leqslant & 2\overline{\alpha}_\varepsilon(\epsilon^\star) \\ & +\Upsilon\left(\tfrac{1-\gamma^\star}{\gamma^\star}+\epsilon^\star,\overline{\alpha}_Y^{-1}\circ\overline{\alpha}_\varepsilon(\epsilon^\star)\right) \end{aligned} \tag{30}$$

We then deduce from the last inequality in (24) that

$$\max\left\{Y^i_{\gamma,\varepsilon}(x),Y^i_{\gamma,\varepsilon}(\upsilon)\right\} \leqslant \tilde{\delta}. \tag{31}$$

The solutions to (6) initialized at x remain in the set $\mathcal{X}$ for all positive times. Indeed, since $x\in\mathcal{X}$ and $Y^i_{\gamma,\varepsilon}(x)\leqslant\Delta$, $Y^i_{\gamma,\varepsilon}(\upsilon)\leqslant Y^i_{\gamma,\varepsilon}(x)\leqslant\Delta$ according to (27). Moreover, (28) and (31) imply that $Y^i_{\gamma,\varepsilon}(\upsilon)\leqslant\tilde{\delta}\leqslant\Delta$ by definition of $\tilde{\delta}$. Hence, in both cases, $Y^i_{\gamma,\varepsilon}(\upsilon)\leqslant\Delta$ which means that $\upsilon\in\mathcal{X}$ as $\left\{z\in\mathbb{R}^{n_x}:Y^i_{\gamma,\varepsilon}(z)\leqslant\Delta\right\}\subset\{z\in\mathbb{R}^{n_x}:\underline{\alpha}_Y(\sigma(z))\leqslant\underline{\alpha}_\varepsilon(\bar{\epsilon})+\Delta\}\subset\mathcal{X}$ 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 $k\in\mathbb{Z}_{\geqslant 0}$ and any solution ϕ to system (6) initialized at x,

$$Y^i_{\gamma,\varepsilon}(\phi(k,x)) \leqslant \max\left\{\beta_Y(Y^i_{\gamma,\varepsilon}(x),k),\tilde{\delta}\right\}. \tag{32}$$

We deduce from item (i) of Theorem 1, since $|\varepsilon|\leqslant\epsilon^\star$, $\underline{\alpha}_Y(\sigma(\phi(k,x)))-\underline{\alpha}_\varepsilon(\epsilon^\star)\leqslant\max\left\{\beta_Y(\overline{\alpha}_Y(\sigma(x))+\overline{\alpha}_\varepsilon(\epsilon^\star),k),\tilde{\delta}\right\}$. Hence, $\sigma(\phi(k,x))\leqslant\underline{\alpha}_Y^{-1}\left(\max\left\{\beta_Y(\overline{\alpha}_Y(\sigma(x))+\overline{\alpha}_\varepsilon(\epsilon^\star),k),\tilde{\delta}\right\}+\underline{\alpha}_\varepsilon(\epsilon^\star)\right)$. Using the property that $\alpha(s_1+s_2)\leqslant\alpha(2s_1)+\alpha(2s_2)$ for any $\alpha\in\mathcal{K}$, $s_1,s_2\in\mathbb{R}_{\geqslant 0}$, see [13, (6)], we derive $\sigma(\phi(k,x))\leqslant\underline{\alpha}_Y^{-1}\Big(\max\Big\{\beta_Y(2\overline{\alpha}_Y(\sigma(x)),k)+\beta_Y(2\overline{\alpha}_\varepsilon(\epsilon^\star),0),\tilde{\delta}\Big\}+\underline{\alpha}_\varepsilon(\epsilon^\star)\Big)$. Thus, by using twice the same property $\sigma(\phi(k,x))\leqslant\underline{\alpha}_Y^{-1}\left(2\max\left\{\beta_Y(2\overline{\alpha}_Y(\sigma(x)),k)+\beta_Y(2\overline{\alpha}_\varepsilon(\epsilon^\star),0),\tilde{\delta}\right\}\right)+\underline{\alpha}_Y^{-1}\left(2\underline{\alpha}_\varepsilon(\epsilon^\star)\right)=\max\left\{\underline{\alpha}_Y^{-1}\big(2\beta_Y(2\overline{\alpha}_Y(\sigma(x)),k)+2\beta_Y(2\overline{\alpha}_\varepsilon(\epsilon^\star),0)\big),\underline{\alpha}_Y^{-1}(2\tilde{\delta})\right\}+\underline{\alpha}_Y^{-1}\left(2\underline{\alpha}_\varepsilon(\epsilon^\star)\right)\leqslant\max\left\{\underline{\alpha}_Y^{-1}\big(4\beta_Y(2\overline{\alpha}_Y(\sigma(x)),k)\big)+\underline{\alpha}_Y^{-1}\big(4\beta_Y(2\overline{\alpha}_\varepsilon(\epsilon^\star),0)\big),\underline{\alpha}_Y^{-1}(2\tilde{\delta})\right\}+\underline{\alpha}_Y^{-1}\left(2\underline{\alpha}_\varepsilon(\epsilon^\star)\right)$. Since $\max\{s_1,s_2\}\leqslant s_1+s_2$ for any $s_1,s_2\in\mathbb{R}_{\geqslant 0}$,

$$\begin{aligned} \sigma(\phi(k,x)) \leqslant & \underline{\alpha}_Y^{-1}\left(4\beta_Y(2\overline{\alpha}_Y(\sigma(x)),k)\right) \\ & +\underline{\alpha}_Y^{-1}\left(4\beta_Y(2\overline{\alpha}_\varepsilon(\epsilon^\star),0)\right)+\underline{\alpha}_Y^{-1}(2\tilde{\delta}) \\ & +\underline{\alpha}_Y^{-1}\left(2\underline{\alpha}_\varepsilon(\epsilon^\star)\right). \end{aligned} \tag{33}$$

[6]It is always possible to find such x in view of Assumption 3 and item (i) of Theorem 1.

By definition of $\tilde{\delta}$ and in view of the second inequality in (24),

$$\sigma(\phi(k,x)) \leqslant \underline{\alpha}_Y^{-1}\left(4\beta_Y(2\overline{\alpha}_Y(\sigma(x)),k)\right)+\delta. \quad (34)$$

This inequality ensures that (12) is satisfied with $\beta(s_1,s_2)=\underline{\alpha}_Y^{-1}\left(4\beta_Y(2\overline{\alpha}_Y(s_1),s_2)\right)$ for any $s_1,s_2\geqslant 0$. ■

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 $\{x\in\mathbb{R}^{n_x}:\underline{\alpha}_Y(\sigma(x))\leqslant\underline{\alpha}_\varepsilon(\tilde{\epsilon})\}$ is included in the interior of $\mathcal{X}$. We define $\Delta\in\mathbb{R}_{>0}$, the biggest positive constant such that $\{x\in\mathbb{R}^{n_x}:\underline{\alpha}_Y(\sigma(x))\leqslant\underline{\alpha}_\varepsilon(\tilde{\epsilon})+\Delta\}\subseteq\mathcal{X}$, which exists in view of the definition of $\tilde{\epsilon}$ above, Assumption 1, and the fact that $\mathcal{X}$ is compact.

Let $i\geqslant i^\star$, $\gamma\in[\gamma^\star,1]$, $\epsilon\in[0,\epsilon^\star]$, $\varepsilon\in\mathcal{B}_{n_\varepsilon}(\epsilon)$, $x\in\mathcal{X}$ such that $Y^i_{\gamma,\varepsilon}(x)\leqslant\Delta$ and $\upsilon\in\widehat{F}_{\gamma,\varepsilon,i}(x)$. In view of (21), (23) and items (ii)-(iii) of Corollary 1,

$$\begin{aligned}Y^i_{\gamma,\varepsilon}(\upsilon)-Y^i_{\gamma,\varepsilon}(x) &\leqslant -a_W\sigma(x)\\ &\quad+\tfrac{1-\gamma}{\gamma}\Big(\bar{a}_V\sigma(x)+a_{\text{stop}}(|\varepsilon|)\sigma(x)\Big)\\ &\quad+a_{\text{stop}}(|\varepsilon|)\sigma(x)\\ &= \Big(-a_W+\frac{1-\gamma}{\gamma}\Big(\bar{a}_V+a_{\text{stop}}(2|\varepsilon|)\Big)\\ &\quad+a_{\text{stop}}(2|\varepsilon|)\Big)\sigma(x).\end{aligned} \quad (35)$$

Since $\gamma\in[\gamma^\star,1]$ and $\epsilon\in[0,\epsilon^\star]$, $-a_W+\frac{1-\gamma}{\gamma}\Big(\bar{a}_V+a_{\text{stop}}(|\varepsilon|)\Big)+a_{\text{stop}}(|\varepsilon|)\leqslant -a_W+\frac{1-\gamma^\star}{\gamma^\star}\Big(\bar{a}_V+a_{\text{stop}}(\epsilon^\star)\Big)+a_{\text{stop}}(\epsilon^\star)<0$. Hence, there exists $\nu>0$ independent of (γ,ϵ) and sufficiently small such that

$$Y^i_{\gamma,\varepsilon}(\upsilon)-Y^i_{\gamma,\varepsilon}(x) \leqslant -\nu\sigma(x). \quad (36)$$

We then apply similar arguments as in the proof of Theorem 2 to obtain the desired result.

When, in addition $\underline{\alpha}_W(s)=\underline{a}_W\cdot s$ for any $s\geqslant 0$, we also have that $\underline{a}_Y\sigma(x)\leqslant Y^i_{\gamma,\varepsilon}(x)\leqslant\overline{a}_Y\sigma(x)$ 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, $\varepsilon_V\in\mathcal{B}_{n_V}(\bar{\epsilon})$ and $i\in\mathbb{Z}_{\geqslant 0}$. In view of item (iv) of Corollary 2, $\ell(x,u)+e_{\widehat{V},i+1}(1,\varepsilon_V,x)\geqslant\ell(x,u)-|\varepsilon_V|\ell(x,0)\geqslant(1-\bar{\epsilon})Q(x)+R(u)$. Recalling that Q and R take non-negative values and $\bar{\epsilon}\in[0,1)$ according to item (iv) of Corollary 2, we derive that $\widehat{V}^i_{1,\varepsilon_V}(x)\geqslant 0$. 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 $x\in\mathcal{X}$, $\varepsilon_V\in\mathcal{B}_{n_V}(\bar{\epsilon})$ and $i\in\mathbb{Z}_{\geqslant 0}$,

$$\begin{aligned}\mathcal{T}\widehat{V}^i_{1,\varepsilon_V}(x) &\leqslant \widehat{V}^i_{1,\varepsilon_V}(x)+2|\varepsilon_V|\breve{V}^i(x)\\ &\leqslant \widehat{V}^i_{1,\varepsilon_V}(x)+\tfrac{2|\varepsilon_V|}{1-|\varepsilon_V|}\overline{V}^0(x),\end{aligned} \quad (37)$$

where $\breve{V}^i(x)$ is defined in [12, Lemma 1] and $\overline{V}^0(x)\leqslant 2\widehat{V}^0(x)$. Therefore, in view of item (vii) of Corollary 2,

$$\mathcal{T}\widehat{V}^i_{1,\varepsilon_V}(x) \leqslant \widehat{V}^i_{i,\varepsilon_V}(x)+\tfrac{2\bar{\epsilon}}{1-\bar{\epsilon}}2dQ(x). \quad (38)$$

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 $i\in\mathbb{Z}_{\geqslant 0}$ and there is no parameter $\varepsilon_{\text{stop}}$. Also, SA3 holds is trivially verified with $e_{\text{stop},i}(1,\varepsilon_V,x)=\mathcal{T}\widehat{V}^i_{\gamma,\varepsilon_V}(x)$ when $x\in\mathbb{R}^{n_x}\backslash\mathcal{X}$. Note that $i^\star=0$ 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 $W=0$, $\overline{\alpha}_W=\alpha_W=a_W\mathbb{I}$ with $a_W=1$ as $\sigma=Q$ and $\ell(x,u)\geqslant\ell(x,0)=Q(x)=\sigma(x)$ 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 $\varepsilon\in\mathcal{B}_{n_V}(\bar{\epsilon})$, 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 $x\in\mathcal{X}$, $i\in\mathbb{Z}_{\geqslant 0}$ and $\varepsilon\in\mathcal{B}_{n_V}(\bar{\epsilon})$, we have $|e_{\widehat{V},i}(1,\varepsilon,x)|\leqslant|\varepsilon|\ell(x,0)=Q(x)-(1-|\varepsilon|)Q(x)\leqslant Q(x)-(1-\bar{\epsilon})Q(x)$. Therefore, (9b) holds with $\underline{\alpha}_W=\underline{a}_W\mathbb{I}$, $\underline{a}_W=1-\bar{\epsilon}$ and $\underline{\alpha}_\varepsilon=0$.

Let $x\in\mathcal{X}$, $\varepsilon\in\mathcal{B}_{n_V}(\bar{\epsilon})$ and $i\geqslant 0$. As $\widehat{V}^i_{1,\varepsilon}(x)\leqslant\overline{V}^i(x)$ according to [12, Theorem 1] and $\overline{V}^i(x)\leqslant\overline{V}^0(x)$ as explained after [12, (36)], $\widehat{V}^i_{1,\varepsilon}(x)\leqslant\overline{V}^0(x)$ and since $\overline{V}^0(x)\leqslant 2dQ(x)$ as shown above, $\widehat{V}^i_{1,\varepsilon}(x)\leqslant 2dQ(x)$. Thus, (10) is verified with $\overline{\alpha}_V=2d\mathbb{I}\in\mathcal{K}_\infty$ and $\overline{\alpha}_\varepsilon=0$: 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.