HAL
open science

Stability of infinite-horizon optimal control with discounted cost

Romain Postoyan, Lucian Busoniu, Dragan Nesic, Jamal Daafouz

To cite this version:

Romain Postoyan, Lucian Busoniu, Dragan Nesic, Jamal Daafouz. Stability of infinite-horizon optimal control with discounted cost. 53rd IEEE Conference on Decision and Control, Dec 2014, Los Angeles, United States. hal-01080084

HAL Id: hal-01080084

https://hal.science/hal-01080084v1

Submitted on 18 Dec 2014

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 of infinite-horizon optimal control with discounted cost

R. Postoyan, L. Buşoniu, D. Nešić, and J. Daafouz

***Abstract*— We investigate the stability of general nonlinear discrete-time systems controlled by an optimal sequence of inputs that minimizes an infinite-horizon discounted cost. We first provide conditions under which a global asymptotic stability property is ensured for the corresponding undiscounted problem. We then show that this property is semiglobally and practically preserved in the discounted case, where the adjustable parameter is the discount factor. We then focus on a scenario where the stage cost is bounded and we explain how our framework applies to guarantee stability in this case. Finally, we provide sufficient conditions, including boundedness of the stage cost, under which the value function, which serves as a Lyapunov function for the analysis, is continuous. As already shown in the literature, the continuity of the Lyapunov function is crucial to ensure some nominal robustness for the closed-loop system.**

I. Introduction

Optimal control selects control inputs so as to minimize a cost incurred during the system operation ([14]). In this paper, we focus on optimal control in discrete time over an infinite horizon, with general nonlinear system dynamics as well as general stage costs. In this general setting, optimal control is a very powerful framework ([3]), able to address decision-making problems not only in control engineering, but also in artificial intelligence, operations research, economy, medicine, etc. In *discounted* optimal control, stage costs are weighted by an exponentially decreasing term γ^k, where $\gamma \in (0,1)$ is the discount factor and k is the time step. The discounted setting is popular in many fields, such as reinforcement learning [24], [25], [6] and dynamic programming [17]. For example, a novel class of predictive optimistic algorithms [18], which provides a very useful relationship between computation invested and near-optimality, only makes sense in the discounted case.

A core practical question is whether the discounted optimal control law stabilizes the system. Some results exist in the economy literature, showing local stability in the continuous-time case *e.g.*, [21], [23], while, in the dynamic programming field, the analysis is usually tailored to the specific cost function considered or the specific algorithm used, see for instance [1], [5], [22]. To the best of our knowledge, the stability properties in the general discounted case are not yet understood. The authors of [15] highlight the difficulties remaining even in the linear case. The main difficulty comes from the impact of the value of γ on the system stability. We will show that even for a simple linear example, γ needs to be sufficiently close to 1 to ensure stability.

Motivated by this insight, we develop a stability analysis for discounted optimal control with general nonlinear dynamics and stage costs, building on the undiscounted problem ($\gamma = 1$) as a nominal case. Contrary to most aforementioned references, we investigate global asymptotic stability properties. Furthermore, we define stability using a generic measure as in [9], which allows one to address the classical equilibrium point stability as a particular case, but also set stability for instance. We first provide sufficient conditions to ensure a global asymptotic stability property for the undiscounted problem, like in [16]. Then, our main result guarantees that this stability property is preserved semiglobally and practically where the adjustable parameter is γ. In other words, for any basin of attraction, the system solutions initialized in this basin will converge to any neighborhood of the target set provided γ is sufficiently close to 1. We only impose standard continuity and compactness conditions on the problem ([3], [10]), together with a controllability property of the system as in [9]. Our analysis is inspired by [9], where the undiscounted finite-horizon case was considered.

Often, the stage cost is taken bounded, see for instance [18], [24], [25], [6] and Chapter 4 of [3]. This has important practical advantages, such as ensuring bounded value functions, as well as a contraction property that guarantees the exponential convergence of important algorithms such as value iteration and policy evaluation. By strengthening the controllability requirement, we explain how our framework applies to a type of bounded stage cost, obtained by saturating an originally unbounded function. Finally, for general continuous bounded stage costs, we show that the Lyapunov (value) function is continuous. This property is essential to endow the system with some nominal robustness, see [8], [12]. Note that when stage costs can grow unbounded, the value function is usually only shown to be lower semicontinuous [10], [7], while *e.g.*, [4] shows continuity for the specific case of concave stage costs.

Next, after introducing some preliminaries in Section II

R. Postoyan and J. Daafouz are with the Université de Lorraine, CRAN, UMR 7039 and the CNRS, CRAN, UMR 7039, France, {romain.postoyan,jamal.daafouz}@univ-lorraine.fr.
Their work was partially supported by the ANR under the grant COMPACS (ANR-13-BS03-0004-02) and the European 7th Framework Network of Excellence 'Highly-complex and networked control systems' (HYCON2) (grant agreement no. 257462).

L. Buşoniu is with the Department of Automation, Technical University of Cluj-Napoca, Romania, lucian.busoniu@aut.utcluj.ro. His work was supported by a grant of the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, project number PNII-RU-TE-2012-3-0040, and by an internal grant of the Technical University of Cluj-Napoca.

D. Nešić is with the Department of Electrical and Electronic Engineering, the University of Melbourne, Parkville, VIC 3010, Australia, dnesic@unimelb.edu.au. His work was supported by the Australian Research Council under the Discovery Projects and Future Fellowship schemes.

and stating the problem in Section III, we provide the nominal, undiscounted analysis in Section IV. Stability with discount is studied in Section V, while Section VI specializes the analysis to bounded stage costs, and Section VII shows the continuity of the Lyapunov function. Conclusions are given in Section VIII.

II. Preliminaries

Let $\mathbb{R} = (-\infty, \infty)$, $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$, $\mathbb{R}_{\geq 0} = [0, \infty)$, $\mathbb{R}_{>0} = (0, \infty)$, $\mathbb{Z}_{\geq 0} = \{0, 1, 2, \ldots\}$, $\mathbb{Z}_{>0} = \{1, 2, \ldots\}$. For $(x, y) \in \mathbb{R}^{n+m}$, (x, y) stands for $[x^{\mathrm{T}}, y^{\mathrm{T}}]^{\mathrm{T}}$. A function $\chi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ 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 it is unbounded. A continuous function $\chi : \mathbb{R}^2_{\geq 0} \longrightarrow \mathbb{R}_{\geq 0}$ 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 $s \in \mathbb{R}_{>0}$, $\chi(s, \cdot)$ is decreasing to zero. We denote by $\mathbb{B}$ the Banach space of bounded real functions on $\mathbb{R}^n$ with the infinity norm $\|v\| = \sup_{x \in \mathbb{R}^n} |v(x)|$.

The definitions below can be found in [20]. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is *lower semicontinuous* at $\bar{x} \in \mathbb{R}^n$ if $\liminf_{x \to \bar{x}} f(x) := \lim_{\delta \searrow 0} [\inf_{x \in \mathcal{B}(\bar{x}, \delta)} f(x)] = f(\bar{x})$, where $\mathcal{B}(\bar{x}, \delta)$ is the closed ball of $\mathbb{R}^n$, centered at $\bar{x}$ of radius $\delta \geq 0$. We say that f is lower semicontinuous on $X \subseteq \mathbb{R}^n$, when it is lower semicontinuous at any $\bar{x} \in X$. Note that when f is continuous, it is also lower semicontinuous.

Let $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ denote a set-valued mapping. The *domain* of S is defined by $\operatorname{dom} S := \{x \in \mathbb{R}^n : S(x) \neq \emptyset\}$. We recall the notions of outer and inner limits of S at $\bar{x} \in \mathbb{R}^n$: $\limsup_{x \to \bar{x}} S(x) := \Big\{u : \exists \{x_n\}_{n \in \mathbb{Z}_{\geq 0}} \exists \{u_n\}_{n \in \mathbb{Z}_{\geq 0}}$ s.t. $x_n \underset{n \to \infty}{\to} \bar{x}$, $u_n \underset{n \to \infty}{\to} u$ with $u_n \in S(x_n)\Big\}$ and $\liminf_{x \to \bar{x}} S(x) := \Big\{u : \forall \{x_n\}_{n \in \mathbb{Z}_{\geq 0}}$ s.t. $x_n \underset{n \to \infty}{\to} \bar{x}$, $\exists \phi \in \mathcal{N}$ $\exists \{u_{\phi(n)}\}_{n \in \mathbb{Z}_{\geq 0}}$ s.t. $u_{\phi(n)} \underset{n \to \infty}{\to} u$ with $u_n \in S(x_n)\Big\}$, where $\mathcal{N}$ is the set of strictly increasing functions from $\mathbb{Z}_{\geq 0}$ to $\mathbb{Z}_{\geq 0}$. When $\limsup_{x \to \bar{x}} S(x) = \liminf_{x \to \bar{x}} S(x)$, we write the limit as $\lim_{x \to \bar{x}} S(x)$. The set-valued mapping S is *continuous* at $\bar{x} \in \mathbb{R}^n$ when $S(x) \to S(\bar{x})$ as $x \to \bar{x}$, and it is continuous on $X \subseteq \mathbb{R}^n$ when it is continuous at any $\bar{x} \in X$. Define the image of a set $\mathcal{V}$ under the mapping S by $S(\mathcal{V}) = \bigcup_{x \in \mathcal{V}} S(x)$. The mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *locally bounded* when for any $\bar{x} \in \mathbb{R}^n$, for some neighborhood $\mathcal{V}$ of $\bar{x}$, the set $S(\mathcal{V}) \subset \mathbb{R}^m$ is bounded.

III. Problem statement

Consider the system

$$x(k+1) \quad = \quad f(x(k), u(k)) \tag{1}$$

with state $x \in \mathbb{R}^n$ and input $u \in \mathcal{U}(x) \subseteq \mathbb{R}^m$, where $\mathcal{U}(x)$ is a non-empty set of admissible inputs associated to state x (as in [10], [11] for instance), and $n, m \in \mathbb{Z}_{>0}$. Let $\mathcal{W} := \{(x, u) : x \in \mathbb{R}^n \text{ and } u \in \mathcal{U}(x)\}$. Define an infinite-length sequence of control inputs $\mathbf{u} = (u_0, u_1, \ldots)$, in which the control input at time $k \in \mathbb{Z}_{\geq 0}$ is given by $u(k) = u_k$. We denote the solution to (1) at the kth-step starting at state x and with the input sequence $\mathbf{u}$ as $\phi(k, x, \mathbf{u}|_k)$, where $\mathbf{u}|_k := (u_0, \ldots, u_{k-1})$ is the truncation of $\mathbf{u}$ to the first $k \in \mathbb{Z}_{>0}$ steps, and we use the convention $\phi(0, x, \mathbf{u}|_0) = x$. We focus on the scenario where the sequence of control inputs is given by the solution to an optimization problem. In particular, we consider the *cost function*

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

where $\ell : \mathcal{W} \to \mathbb{R}$ is the *stage cost* and $\gamma \in (0, 1)$ is the *discount factor*. The objective is to minimize the cost J. When $\inf_{\mathbf{u}} J(x, \mathbf{u})$ is well-defined (which is assumed to be the case in this study), we call it the *optimal value function* and we denote it

$$V(x) \quad := \quad \inf_{\mathbf{u}} J(x, \mathbf{u}). \tag{3}$$

Further, when the infimum in (3) is achieved by an infinite-length input sequence, we denote this sequence $\mathbf{u}^*(x)$ and call it the *optimal solution* at state x. Thus, $V(x) = J(x, \mathbf{u}^*(x))$.

The objective of this study is to infer about the stability property of system (1) when controlled by the optimal sequence $\mathbf{u}^*(x)$ (which may be non-unique), when the latter exists. Our approach consists in considering, as a nominal case, the corresponding undiscounted problem, *i.e.* when the cost function is given by (2) with $\gamma = 1$

$$\overline{J}(x, \mathbf{u}) \quad := \quad \sum_{k=0}^{\infty} \ell(\phi(k, x, \mathbf{u}|_k), u_k), \tag{4}$$

and the optimal value function is $\overline{V}(x) := \inf_{\mathbf{u}} \overline{J}(x, \mathbf{u})$. In this undiscounted setting, we can guarantee the stability of the system with an optimal sequence of inputs $\overline{\mathbf{u}}^*(x)$ (which will be ensured to exist), defined as a sequence so that $\overline{V}(x) = \overline{J}(x, \overline{\mathbf{u}}^*(x))$. We then study under which conditions it is possible to (approximately) maintain this stability property in the discounted case, when $\gamma \in (0, 1)$.

Our approach is motivated by the study of a simple example which shows that γ indeed needs to be appropriately selected to preserve stability.

Example 1: Consider the scalar system $x(k+1) = 2x(k) + u(k)$ and the discounted quadratic cost $J(x, \mathbf{u}) = \sum_{k=0}^{\infty} \gamma^k \left(x(k)^2 + u(k)^2\right)$, where $\gamma \in (0, 1)$ and $\mathbf{u}$ is an infinite-length sequence of inputs. The optimal solution is given by the feedback law $u^* = Kx$ with $K = -2\left(1 + 2\left(5\gamma - 1 + \sqrt{(5\gamma - 1)^2 + 4\gamma}\right)^{-1}\right)^{-1}$ (see Section 4.2 in [2]). The origin of the closed-loop system is globally exponentially stable if and only if $2 + K \in (-1, 1)$, which is equivalent to $\gamma \in (\gamma^\star, 1)$ where $\gamma^\star = \frac{1}{3}$. Hence, γ needs to be large enough, otherwise the optimal feedback law does not stabilize the system. $\square$

IV. Preliminaries on the undiscounted problem

We make the following assumption on the undiscounted problem, which guarantees the existence of an optimal solution according to [11].

Assumption 1: The following holds.

(i) f is continuous on $\mathcal{W}$ and ℓ is lower semicontinuous and nonnegative on $\mathcal{W}$.

(ii) For any compact set $\mathcal{C}$, the set $\{(x,u) \in \mathcal{W} : x \in \mathcal{C}\}$ is compact.

(iii) For any $x \in \mathbb{R}^n$, there exists an admissible sequence of inputs $\mathbf{u}$ such that $\overline{J}(x,\mathbf{u})$ is bounded. □

Assumption 1 corresponds to the conditions of Theorem 1 in [11]. Item (ii) of Assumption 1 is satisfied *e.g.*, when $\mathcal{U}(x) = U$ for any $x \in \mathbb{R}^n$ with U compact; more general conditions are given in Theorem 2 in [11]. Item (iii) of Assumption 1 is verified for instance when ℓ satisfies a general exponential controllability property introduced in Definition 3 in [9], as will be formalized in the lemma below. This property is stated with respect to a generic continuous, positive definite function $\sigma : \mathbb{R} \to \mathbb{R}_{\geq 0}$, which will serve later as a state measure (as in [9]). The function σ can be defined *e.g.*, as $|\cdot|$, $|\cdot|^2$ or by the distance to a set.

Lemma 1: Consider system (1) and suppose that ℓ is *globally exponentially controllable to zero with respect to* $\sigma : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, *i.e.* there exist $M, \lambda > 0$ such that for any $x \in \mathbb{R}^n$ there exists an admissible infinite-length control input sequence $\mathbf{u}$ such that $\ell(\phi(k,x,\mathbf{u}|_k),u_k) \leq M\sigma(x)e^{-\lambda k}$ for any $k \in \mathbb{Z}_{>0}$. Then item (iii) of Assumption 1 holds. □

Proof. Let $x \in \mathbb{R}^n$ and take the sequence $\mathbf{u}$ that satisfies the condition of Lemma 1. It then holds that, for any $N \in \mathbb{Z}_{>0}$,

$$\begin{aligned} \sum_{k=0}^{N} \ell(\phi(k,x,\mathbf{u}|_k),u_k) &\leq \sum_{k=0}^{N} M\sigma(x)e^{-\lambda k} \\ &\leq \frac{M\sigma(x)}{1-e^{-\lambda}}. \end{aligned} \tag{5}$$

Since the inequalities above hold for any $N \in \mathbb{Z}_{>0}$, they also hold when $N \to \infty$, and $\overline{J}(x,\mathbf{u}) = \sum_{k=0}^{\infty} \ell(\phi(k,x,\mathbf{u}|_k),u_k) \leq \frac{M\sigma(x)}{1-e^{-\lambda}}$, which implies the satisfaction of item (iii) of Assumption 1. ■

The conditions below come from [9] where the stability of model predictive control problems with finite-horizon undiscounted cost is investigated. These same conditions will allow us to guarantee a global asymptotic stability property for system (1) when the cost is infinite-horizon and undiscounted, similarly to [16].

Assumption 2: Let $\sigma : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ be continuous and positive definite. There exist $\overline{\rho}, \overline{\alpha} \in \mathcal{K}_\infty$ such that the following holds for any $(x,u) \in \mathcal{W}$

$$\overline{V}(x) \leq \overline{\alpha}(\sigma(x)) \quad \text{and} \quad \ell(x,u) \geq \overline{\rho}(\sigma(x)). \tag{6}$$

□

The first inequality in (6) is verified *e.g.*, when ℓ is globally exponentially controllable to zero with respect to $\sigma : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$. Indeed, in that case, we have that $\overline{V}(x) \leq \overline{\alpha}(\sigma(x))$ with $\overline{\alpha}(\sigma(x)) = \frac{M\sigma(x)}{1-e^{-\lambda}}$ in view of the proof of Lemma 1. The second inequality in (6) is a detectability property of σ from ℓ (see Definition 1 in [9]), which is satisfied for instance when ℓ is quadratic in its first argument and $\sigma(\cdot) = |\cdot|$. This detectability property can be relaxed as in SA3 in [9]; we will address this extension in future work.

The theorem below formalizes the existence of an achievable sequence which minimizes the cost function in (4) (item (a)) and states a global asymptotic stability property for system (1) when this sequence is applied (item (b)).

Theorem 1: Consider system (1), the cost function $\overline{J}$ in (4), and suppose Assumptions 1-2 hold. Then the following holds.

(a) For any $x \in \mathbb{R}^n$ there exists an optimal sequence of inputs $\overline{\mathbf{u}}^*(x)$.

(b) There exists $\beta \in \mathcal{KL}$ such that for any $x \in \mathbb{R}^n$ and $k \in \mathbb{Z}_{\geq 0}$, $\sigma(\phi(k,x,\overline{\mathbf{u}}^*(x)|_k)) \leq \beta(\sigma(x),k)$. □

Proof. Assumption 1 allows us to directly apply Theorem 1 in [11] which ensures the satisfaction of item (a) of Theorem 1. Let $x \in \mathbb{R}^n$ and consider $\overline{\mathbf{u}}^*(x)$. According to the Bellman equation,

$$\overline{V}(x) = \ell(x,\overline{u}_0^*) + \overline{V}(\phi(1,x,\overline{\mathbf{u}}^*(x)|_1)). \tag{7}$$

Hence, in view of Assumption 2,

$$\begin{aligned} \overline{V}(\phi(1,x,\overline{\mathbf{u}}^*(x)|_1)) - \overline{V}(x) &= -\ell(x,\overline{u}_0^*) \leq -\overline{\rho}(\sigma(x)) \\ &\leq -\overline{\rho} \circ \overline{\alpha}^{-1}(\overline{V}(x)). \end{aligned} \tag{8}$$

We obtain by induction, for any $k \in \mathbb{Z}_{\geq 0}$,

$$\begin{aligned} &\overline{V}(\phi(k+1,x,\overline{\mathbf{u}}^*(x)|_{k+1})) - \overline{V}(\phi(k,x,\overline{\mathbf{u}}^*(x)|_k)) \\ &\qquad \leq -\overline{\rho} \circ \overline{\alpha}^{-1}(\overline{V}(\phi(k,x,\overline{\mathbf{u}}^*(x)|_k))). \end{aligned} \tag{9}$$

As a consequence, according to [19], there exists $\overline{\beta} \in \mathcal{KL}$ such that for any $k \in \mathbb{Z}_{>0}$, $\overline{V}(\phi(k,x,\overline{\mathbf{u}}^*(x)|_k)) \leq \overline{\beta}(\overline{V}(x),k)$. Using the fact that $\overline{\rho}(\sigma(x)) \leq \ell(x,u_0^*) \leq \overline{V}(x) \leq \overline{\alpha}(\sigma(x))$ in view of Assumption 2, we deduce that item (b) of Theorem 1 holds with $\beta(s,k) = \overline{\rho}^{-1} \circ (\overline{\beta}(\overline{\alpha}(s),k))$ for any $s,k \in \mathbb{R}_{\geq 0}$. ■

Item (b) of Theorem 3 is equivalent to the usual (ε,δ)-definition of global asymptotic stability and is standard in the nonlinear systems literature. It means that the state remains within a region that depends on the magnitude of the initial state, but shrinks with time.

V. Analysis of the discounted problem

Even in the discounted case, Assumption 1 ensures the existence of an optimal sequence of inputs for any initial condition, according to the proposition below.

Proposition 1: Consider system (1), the cost function J in (2), and suppose Assumption 1 holds. For any $x \in \mathbb{R}^n$ there exists an optimal sequence of inputs $\mathbf{u}^*(x)$. □

Proof. Let $\gamma \in (0,1)$. As before, conditions a)-d) of Theorem 1 in [11] are verified in view of items (i)-(ii) of Assumption 1. Let $x \in \mathbb{R}^n$. According to item (iii) of Assumption 1, there exists $\overline{\mathbf{u}}$ such that $\overline{J}(x,\overline{\mathbf{u}})$ is finite. Hence, since $\gamma \in (0,1)$ and ℓ is nonnegative in view of item (i) of Assumption 1, for any $N \in \mathbb{Z}_{>0}$

$$\begin{aligned} 0 &\leq \sum_{k=0}^{N} \gamma^k \ell(\phi(k,x,\overline{\mathbf{u}}|_k),\overline{u}_k) \\ &\leq \sum_{k=0}^{N} \ell(\phi(k,x,\overline{\mathbf{u}}|_k),\overline{u}_k) \leq \overline{J}(x,\overline{\mathbf{u}}). \end{aligned} \tag{10}$$

Since these inequalities hold for any N, they also hold as $N \to \infty$. We deduce that $J(x, \overline{\mathbf{u}}) \in [0, \overline{J}(x, \overline{\mathbf{u}})]$. Hence, $J(x, \overline{\mathbf{u}})$ is bounded and condition e) of Theorem 1 in [11] holds. We apply Theorem 1 in [11] to conclude that the infimum in (3) is well-defined and achievable for any $x \in \mathbb{R}^n$. ■

The theorem below gives Lyapunov-based properties, from which we will later deduce a stability property for system (1) when applying an optimal sequence of inputs.

Theorem 2: Consider system (1) and the cost function J in (2). The following holds under Assumptions 1-2.

(a) For any $x \in \mathbb{R}^n$ and any $\gamma \in (0,1)$,

$$\overline{\rho}(\sigma(x)) \quad \leq \quad V(x) \quad \leq \quad \overline{\alpha}(\sigma(x)). \tag{11}$$

(b) For any $\gamma_m \in (0,1)$, there exists $\Upsilon_{\gamma_m} \in \mathcal{KL}$ such that for any $x \in \mathbb{R}^n$, $\gamma \in (\gamma_m, 1)$, and $k \in \mathbb{Z}_{\geq 0}$

$$\begin{aligned} V(x(k+1)) - V(x(k)) \quad \leq \quad & -\overline{\rho}(\sigma(x(k))) \\ & +\Upsilon_{\gamma_m}(\sigma(x(k)), \tfrac{\gamma}{1-\gamma}), \end{aligned} \tag{12}$$

where $x(k) = \phi(k, x, \mathbf{u}^*(x)|_k)$ with $x \in \mathbb{R}^n$ and $\mathbf{u}^*(x)$ the associated optimal sequence. □

Proof. Let $\gamma \in (0,1)$, $x \in \mathbb{R}^n$, $V(x)$ is well-defined and there exists an optimal sequence $\mathbf{u}^*(x) = (u_0^*, u_1^*, \ldots)$ according to Proposition 1. Since ℓ is nonnegative according to item (i) of Assumption 1 and (6) holds,

$$V(x) \quad \geq \quad \ell(x, u_0^*) \quad \geq \quad \overline{\rho}(\sigma(x)). \tag{13}$$

On the other hand, since $\gamma \in (0,1)$,

$$\begin{aligned} V(x) \quad &= \quad \inf_{\mathbf{u}} \sum_{k=0}^{\infty} \gamma^k \ell(\phi(k, x, \mathbf{u}|_k), u_k) \\ &\leq \quad \inf_{\mathbf{u}} \sum_{k=0}^{\infty} \ell(\phi(k, x, \mathbf{u}|_k), u_k) = \overline{V}(x) \end{aligned} \tag{14}$$

and according to Assumption 2,

$$V(x) \quad \leq \quad \overline{V}(x) \quad \leq \quad \overline{\alpha}(\sigma(x)). \tag{15}$$

From (13) and (15), we have that item (a) of Theorem 2 holds.

We now prove that item (b) of Theorem 2 holds. Let $\gamma_m \in (0,1)$, $\gamma \in (\gamma_m, 1)$, $x \in \mathbb{R}^n$ and u_0^* be defined as above. We denote $\phi(1, x, \mathbf{u}^*(x)|_1) = f(x, u_0^*(x))$ as x^+ for the sake of convenience. According to the Bellman equation,

$$V(x) \quad = \quad \ell(x, u_0^*) + \gamma V(x^+), \tag{16}$$

therefore

$$V(x^+) - V(x) \quad = \quad -\ell(x, u_0^*) + (1-\gamma) V(x^+). \tag{17}$$

Let $\overline{x}^+ := f(x, \overline{u}_0^*(x))$ where $\overline{\mathbf{u}}^*(x) = (\overline{u}_0^*(x), \overline{u}_1^*(x), \ldots)$ is an optimal sequence of inputs for the undiscounted cost. We note that $V(x) = \inf_{\mathbf{u}} J(x, \mathbf{u}) \leq J(x, (\overline{u}_0^*(x), \mathbf{u}^*(\overline{x}^+)) = \ell(x, \overline{u}_0^*) + \gamma V(\overline{x}^+)$. From (16), $\gamma V(x^+) \leq \ell(x, u_0^*) + \gamma V(x^+) \leq \ell(x, \overline{u}_0^*) + \gamma V(\overline{x}^+)$. Hence $\gamma V(x^+) \leq \overline{\alpha}(\sigma(x)) + \gamma \overline{V}(\overline{x}^+)$ from (15) and since $\ell(x, \overline{u}_0^*) \leq \overline{V}(x) \leq \overline{\alpha}(\sigma(x))$. From (9), it follows directly that $\overline{V}(\overline{x}^+) \leq (\mathbb{I} - \overline{\rho} \circ \overline{\alpha}^{-1})(\overline{V}(x))$. Using $\overline{V}(x) \leq \overline{\alpha}(\sigma(x))$ from Assumption 2, we deduce that $\overline{V}(\overline{x}^+) \leq (\overline{\alpha} - \overline{\rho})(\sigma(x))$. Consequently, $\gamma V(x^+) \leq \overline{\alpha}(\sigma(x)) + \gamma(\overline{\alpha} - \overline{\rho})(\sigma(x))$. Then, by using this expression in (17), we have $V(x^+) - V(x) \leq -\overline{\rho}(\sigma(x)) + (1-\gamma)\big(\gamma^{-1}\overline{\alpha}(\sigma(x)) + (\overline{\alpha} - \overline{\rho})(\sigma(x))\big) \leq -\overline{\rho}(\sigma(x)) + (\gamma^{-1} - 1)\big(2\overline{\alpha}(\sigma(x)) - \overline{\rho}(\sigma(x))\big)$, where $\gamma^{-1} \geq 1$ was used to obtain the last inequality. We assume without loss of generality that $2\overline{\alpha} - \overline{\rho} \in \mathcal{K}$ (we can always upper-bound $2\overline{\alpha} - \overline{\rho}$ by a class-$\mathcal{K}$ function if this is not the case by following Lemma 4.3 in [13] as this function is zero at zero, nonnegative and continuous). We also note that $0 \leq \gamma^{-1} - 1 = \min\{\gamma^{-1} - 1, \gamma_m^{-1} - 1\}$ since $\gamma \in (\gamma_m, 1)$. By induction on the time, we deduce that item (b) of Theorem 2 holds with $\Upsilon_{\gamma_m}(s,t) = \min\{t^{-1}, \gamma_m^{-1} - 1\}\big(2\overline{\alpha}(s) - \overline{\rho}(s)\big)$ for any $s, t \geq 0$. ■

The following theorem means that, if the optimal solution for the undiscounted cost globally asymptotically stabilizes the system (in the sense that item (b) of Theorem 1 holds), then the discounted optimal solution will also stabilize the system but semiglobally and practically, where the adjustable parameter is γ.

Theorem 3: Consider system (1), the cost function J in (2), and suppose Assumptions 1-2 hold. There exists $\beta \in \mathcal{KL}$ such that for any $\delta, \Delta > 0$, there exists $\gamma^\star \in (0,1)$ such that for any $\gamma \in (\gamma^\star, 1)$ and $x \in \{z \in \mathbb{R}^n : \sigma(z) \leq \Delta\}$, the solution ϕ to system (1) with initial condition x and the optimal sequence of inputs $\mathbf{u}^*(x)$ satisfies

$$\sigma(\phi(k, x, \mathbf{u}^*(x)|_k)) \quad \leq \quad \max\{\beta(\sigma(x), k), \delta\} \quad \forall k \in \mathbb{Z}_{\geq 0}. \tag{18}$$

□

Proof. We want to use similar arguments as in the proof of Corollary 1 in [9] to obtain the desired result. For that purpose, we need to consider an autonomous system. For any $x \in \mathbb{R}^n$, take the optimal sequence $\mathbf{u}^*(x)$. We write system (1) with input sequence $\mathbf{u}^*(x)$ as $x(k+1) = \tilde{f}(x(k), k)$. Adding the time k as an extra variable, we obtain the autonomous system $\tilde{x} = F(\tilde{x})$ where $\tilde{x} = (x, k)$ and $F(\tilde{x}) = (\tilde{f}(x, k), k+1)$. We define $\tilde{V}(\tilde{x}) := V(x)$ and $\tilde{\sigma}(\tilde{x}) := \sigma(x)$ for any $\tilde{x} = (x, k)$. From Theorem 2, it holds that, for any $\gamma_m \in (0,1)$, $\gamma \in (\gamma_m, 1)$, $\tilde{x} \in \mathbb{R}^{n+1}$,

$$\overline{\rho}(\tilde{\sigma}(\tilde{x})) \quad \leq \quad \tilde{V}(\tilde{x}) \quad \leq \quad \overline{\alpha}(\tilde{\sigma}(\tilde{x})). \tag{19}$$

and, for any $k \in \mathbb{Z}_{\geq 0}$,

$$\begin{aligned} \tilde{V}(\tilde{\phi}(k+1, \tilde{x})) - V(\tilde{\phi}(k, \tilde{x})) \leq & -\overline{\rho}(\tilde{\sigma}(\tilde{\phi}(k, \tilde{x}))) \\ & +\Upsilon_{\gamma_m}(\tilde{\sigma}(\tilde{\phi}(k, \tilde{x})), \tfrac{\gamma}{1-\gamma}), \end{aligned} \tag{20}$$

where $\tilde{\phi}(k, \tilde{x})$ is the solution to system $\tilde{x}(k+1) = F(\tilde{x})$ at time k, with initial condition $\tilde{x}$. Like in the proof of Corollary 1 in [9], we then derive that for any $\delta, \Delta > 0$, there exists $\gamma^\star \in (\gamma_m, 1)$ such that for any $\gamma \in (\gamma^\star, 1)$ and $\tilde{x} \in \{\tilde{z} \in \mathbb{R}^{n+1} : \tilde{\sigma}(\tilde{z}) \leq \Delta\}$, the solution $\tilde{\phi}$ to system $\tilde{x}(k+1) = F(\tilde{x})$ with initial condition $\tilde{x}$ satisfies

$$\tilde{\sigma}(\tilde{\phi}(k, \tilde{x})) \quad \leq \quad \max\{\beta(\tilde{\sigma}(\tilde{x}), k), \delta\} \qquad \forall k \in \mathbb{Z}_{\geq 0}. \tag{21}$$

Since $\tilde{\sigma}(\tilde{x}) = \sigma(x)$ for any $\tilde{x} = (x, k)$, we deduce from (21) that (18) holds. ■

This theorem is one of the main results of the paper. It is interesting to note the analogy with the results in [9] where the finite horizon N takes the place of $\frac{\gamma}{1-\gamma}$ in the infinite-horizon discounted problem. Informally, quantity $\frac{\gamma}{1-\gamma}$ can be thought of as an 'effective horizon' of the discounted problem. Thus, while [9] shows stability for horizons N greater than some lower bound $N^\star$, we show it for $\gamma > \gamma^\star$, and thus for effective horizons larger than $\frac{\gamma^\star}{1-\gamma^\star}$.

The value function V is used in this section as a Lyapunov function to analyse stability. In practice, it is essential to work with a Lyapunov function which is continuous in order to guarantee some nominal robustness for the system (see [12]). The assumptions made so far do not a priori allow us to assert that V is continuous. In Section VII, we will provide conditions which will allow us to prove the continuity of V. These conditions include the boundedness of the stage cost, so we first explain, in Section VI below, how our stability guarantees can be applied to the case of bounded stage costs, which is also of practical importance in its own right.

VI. Bounded stage cost

In many cases, the stage cost is taken bounded, see for instance Chapter 4 in [3] as well as [6], [18] to mention a few. We explain how the results of the previous sections can be applied when the stage cost ℓ is obtained by saturating a stage cost ℓ' which verifies the assumptions of Section IV.

We thus consider an a priori unbounded stage cost $\ell' : \mathcal{W} \to \mathbb{R}$ and a state measure $\sigma' : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$. We denote the undiscounted cost function with stage cost ℓ' as $\overline{J}'$. We focus on the scenario where Assumption 1-2 hold for system (1) with cost function $\overline{J}'$. In addition, we assume that the properties below are verified.

Assumption 3:

(i) There exist $M, \lambda > 0$ such that for any $x \in \mathbb{R}^n$ there exists an infinite-length control input sequence $\mathbf{u}'$ such that $\ell'(\phi(k, x, \mathbf{u}'|_k), u'_k) \leq M\sigma'(x)e^{-\lambda k}$ for any $k \in \mathbb{Z}_{>0}$.

(ii) For any $\mu > 0$ there exists $k_\mu \in \mathbb{Z}_{\geq 0}$ such that for any $x \in \mathbb{R}^n$ there exists an infinite-length control input sequence $\mathbf{u}^\mu$ such that $\ell'(\phi(k, x, \mathbf{u}^\mu|_k), u^\mu_k) \leq \mu$ for any $k \in \mathbb{Z}_{\geq 0}$ with $k \geq k_\mu$. □

Item (i) of Assumption 3 corresponds to the exponential controllability property considered in Lemma 1. Item (b) of Proposition 2 is an uniform finite-time controllability property of ℓ'. It is likely possible to relax this condition to hold semiglobally, *i.e.* to make k_μ depend on the ball of initial conditions; this extension is left for future work.

We now saturate ℓ' by an arbitrary constant $\nu > 0$. In that way, the stage cost becomes

$$\ell(x, u) \quad := \quad \min\{\ell'(x, u), \nu\} \qquad \forall (x, u) \in \mathcal{W}. \tag{22}$$

In this case, we no longer consider the state measure σ' but σ which is defined as

$$\sigma(x) \quad := \quad \min\{\rho'(\sigma'(x)), \nu\} \qquad \forall x \in \mathbb{R}^n, \tag{23}$$

where $\rho' \in \mathcal{K}_\infty$ is such that $\rho'(\sigma'(x)) \leq \ell'(x, u)$ for any $(x, u) \in \mathcal{W}$ (such a function exists since Assumption 2 is assumed to hold when the cost function is given by $\overline{J}'$). We denote the undiscounted cost function with stage cost ℓ as $\overline{J}$. The proposition below states that Assumptions 1-2 hold for system (1) with cost function $\overline{J}$. Hence, the results of Section IV apply in this case with the state measure σ in (23). The proof is omitted due to space limitations.

Proposition 2: Suppose the following holds.

(a) Assumptions 1-2 are verified for system (1) with cost function $\overline{J}'$.

(b) Assumption 3 is verified.

Then Assumptions 1-2 are guaranteed for system (1) with the cost function $\overline{J}$ and the conclusions of Theorems 1-3 hold with σ defined in (23). □

In view of Proposition 2, the state measure σ in (23) is ensured to semiglobally practically converge to the origin in the discounted case (since the conclusions of Theorem 3 apply). As a consequence, so does $\rho'(\sigma')$ since σ becomes equal to $\rho'(\sigma')$ as it approaches the origin, when δ is selected smaller than ν in Theorem 3. We therefore see that the semiglobal practical convergence of σ', equivalently of $\rho'(\sigma')$ as $\rho' \in \mathcal{K}_\infty$, is preserved when saturating the stage cost function ℓ'. That means for instance that, when $\sigma'(\cdot) = |\cdot|$, the system solution converges to a neighborhood of the origin, which can be rendered as small as desired by decreasing δ (*i.e.* by increasing γ), when the stage cost is saturated as in (22). The main difference appears for the stability. Indeed, to saturate ℓ' does not allow us to bound σ' (equivalently $\rho'(\sigma')$) for all time by a class-$\mathcal{K}_\infty$ function of its initial value. However, we know that σ' is bounded as it converges to a neighborhood of the origin as time grows, which may be sufficient in many problems.

VII. Continuity of the value function V

The purpose of this section is to provide conditions under which V is ensured to be continuous. In addition, we will guarantee the existence of a stationary, state-feedback optimal control law.

Assumption 4: The stage cost function ℓ is continuous on $\mathcal{W}$ and bounded, *i.e.* there exists $\nu > 0$ such that $\ell(x, u) \leq \nu$ for any $(x, u) \in \mathcal{W}$. Furthermore, the set-valued mapping $x \mapsto \mathcal{U}(x)$ is continuous and locally bounded. □

We are ready to state the main result of this section.

Proposition 3: Consider system (1), the cost function J in (2) with any $\gamma \in (0, 1)$, and suppose Assumptions 1 and 4 are satisfied. Then the following holds.

(a) The optimal value function V is continuous on $\mathbb{R}^n$.

(b) There exists an optimal solution which is given by a stationary, state-feedback law, *i.e.* there exists $\kappa : \mathbb{R}^n \to \mathbb{R}^m$ such that any $x \in \mathbb{R}^n$, $\kappa(x) \in \mathcal{U}(x)$ and the optimal sequence $\mathbf{u}^*(x) = (u^*_0, u^*_1, \dots)$ can be constructed as follows: $u^*_k = \kappa(\phi(k, x, \mathbf{u}^*(x)|_k))$, for any $k \in \mathbb{Z}_{\geq 0}$. □

Proof. We begin by observing that the boundedness of the stage cost from Assumption 4, its nonnegativity from item

(i) of Assumption 1, and the existence of an optimal solution from Proposition 1, ensure the satisfaction of Assumption C of [3] (the contraction assumption) by Proposition 4.10 of [3]. This assumption in turn allows us to apply Propositions 4.1, 4.2, and 4.3 of [3], which are essential to this proof.

We define the Bellman mapping $T : \mathbb{B} \to \mathbb{B}$, where $[T(v)](x) := \min_{u \in \mathcal{U}(x)} [\ell(x,u) + \gamma v(f(x,u))]$. Item (c) of Proposition 4.1 in [3] shows that T is a contraction of modulus γ, *i.e.* that $\|T(v) - T(v')\| \leq \gamma \|v - v'\|$ for any $v, v' \in \mathbb{B}$, recalling also that $\|\cdot\|$ denotes the infinity norm. Due to item (a) of Proposition 4.2 in [3], we obtain that the optimal value function V is the unique fixed point of this mapping, *i.e.* $V = T(V)$. By item (c) of Proposition 4.2 in [3], the repeated application of T converges to V, that is, the sequence $v_0 = 0$, $v_{n+1} = T(v_n)$ satisfies $\lim_{n\to\infty} \|V - v_n\| = 0$, with $n \in \mathbb{Z}_{\geq 0}$.

Using these facts, the following standard result also holds, by an immediate induction $\|V - v_n\| \leq \gamma^n \|V - v_0\| = \gamma^n \|V\|$, so that the convergence of v_n, $n \in \mathbb{Z}_{\geq 0}$, to V is *uniform*. We now prove by induction that v_n is continuous for any $n \in \mathbb{Z}_{\geq 0}$. This assertion is true for $n = 0$ since $v_0 = 0$, suppose then that it holds for some $n \in \mathbb{Z}_{\geq 0}$. We have $v_{n+1} : x \mapsto \min_{u \in \mathcal{U}(x)} [\ell(x,u) + \gamma v_n(f(x,u))]$. Since $\mathcal{U}$ is continuous, locally bounded (in view of Assumption 4), $\mathrm{dom}\,\mathcal{U} = \mathbb{R}^n$ (by assumption), and $\ell + \gamma v_n$ is continuous, we apply Lemma 2 given in the Appendix to conclude that v_{n+1} is continuous. Consequently, v_n is continuous for any $n \in \mathbb{Z}_{\geq 0}$. Hence, we have a sequence of continuous functions that converges uniformly to V, which allows us to reach our final conclusion that V is continuous, according to the uniform limit theorem. ■

VIII. Conclusions

We have shown that the solutions of a discounted optimal control problem preserves in a certain sense the stability properties of the undiscounted solution, by adjusting the discount factor. Further, the stability guarantees remain valid for a type of bounded stage costs, and the Lyapunov value function can be shown to be continuous when the stage costs are bounded, under mild assumptions.

An immediate line of future work is relaxing the detectability condition on the stage cost in (6) along the lines in [9], and the finite-time controllability condition as explained after Assumption 3. A second direction is motivated by the fact that practical algorithms rarely achieve exactly the optimal solution, ensuring instead only some bounded near-optimality. Further, methods originating in artificial intelligence often require the inputs to be discretized, see [24], [25], [6]. It would therefore be interesting to take into account these sources of errors, aiming to ensure stability for a wide class of practical algorithms.

References

[1] S.N. Balakrishnan, J. Ding, and F.L. Lewis. Issues on stability of ADP feedback controllers for dynamical systems. *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, 38(4):913–917, 2008.

[2] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume II. Athena Scientific, 4th edition, 2012.

[3] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control: The Discrete Time Case*. Academic Press, 1978.

[4] L. Blume, D. Easley, and M. O'Hara. Characterization of optimal plans for stochastic dynamic programs. *Journal of Economic Theory*, 28(2):221–234, 1982.

[5] C.I. Boussios, M.A. Dahleh, and J.N. Tsitsiklis. Semiglobal nonlinear stabilization via approximate policy iteration. In *American Control Conference, Arlington, U.S.A.*, volume 6, pages 4675–4680, 2001.

[6] L. Buşoniu, R. Babuška, B. De Schutter, and D. Ernst. *Reinforcement Learning and Dynamic Programming Using Function Approximators*. Automation and Control Engineering. Taylor & Francis CRC Press, 2010.

[7] R. Cavazos-Cadena and F. Salem-Silva. The discounted method and equivalence of average criteria for risk-sensitive Markov decision processes on Borel spaces. *Applied Mathematics and Optimization*, 61(2):167–190, 2010.

[8] G. Grimm, M.J. Messina, S.E. Tuna, and A.R. Teel. Examples when nonlinear model predictive control is nonrobust. *Automatica*, 40(10):1729–1738, 2004.

[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] O. Hernández-Lerma and J.-B. Lasserre. *Further topics on discrete-time Markov control processes*. Springer, 1999.

[11] S.S. Keerthi and E.G. Gilbert. An existence theorem for discrete-time infinite-horizon optimal control problems. *IEEE Transactions on Automatic Control*, 30(9):907–909, 1985.

[12] 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.

[13] H.K. Khalil. *Nonlinear systems*. Prentice-Hall, Englewood Cliffs, New Jersey, U.S.A., 3rd edition, 2002.

[14] D. E. Kirk. *Optimal Control Theory: An Introduction*. Dover Publications, 2004.

[15] B. Kiumarsi, F. L. Lewis, H. Modares, A. Karimpour, and M.-B. Naghibi-Sistani. Reinforcement Q-learning for optimal tracking control of linear discrete-time systems with unknown dynamics. *Automatica*, 50(4):1167–1175, 2014.

[16] G. Kreisselmeier and T. Birkhölzer. Numerical nonlinear regulator design. *IEEE Transactions on Automatic Control*, 39(1):33–46, 1994.

[17] F. Lewis and D. Liu, editors. *Reinforcement Learning and Adaptive Dynamic Programming for Feedback Control*. Wiley, 2012.

[18] 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, 2013.

[19] D. Nešić and A.R. Teel. A framework for stabilization of nonlinear sampled-data systems based on their approximate discrete-time. *IEEE Transactions on Automatic Control*, 49:1103–1122, 2004.

[20] R.T. Rockafellar and R.J.-B. Wets. *Variational analysis*, volume 317. Springer, 3rd edition, 1998.

[21] A. Rodriguez. On the local stability of the solution to optimal control problems. *Journal of Economic Dynamics and Control*, 28(12):2475–2484, 2004.

[22] Y. Sokolov, R. Kozma, L. D. Werbos, and P. J. Werbos. Complete stability analysis of a heuristic ADP control design. *arXiv*, 2013.

[23] G. Sorger. Local stability of stationary states in discounted optimal control systems. *Journal of Optimization Theory and Applications*, 72(1):143–162, 1992.

[24] R.S. Sutton and A.G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[25] C. Szepesvári. *Algorithms for Reinforcement Learning*. Morgan & Claypool Publishers, 2010.

Appendix

The following lemma is needed to prove Proposition 3. Its proof is omitted due to space constraints.

Lemma 2: Let $\varphi : \mathbb{R}^{n+m} \to \mathbb{R}_{\geq 0}$ be continuous and the set-valued mapping $\mathcal{U} : \mathrm{dom}\,\mathcal{U} \rightrightarrows \mathbb{R}^m$ be continuous and locally bounded and $\mathrm{dom}\,\mathcal{U} = \mathbb{R}^n$. Then $W : x \mapsto \inf_{u \in \mathcal{U}(x)} \varphi(x,u)$ is continuous. □